8-2017

Examining the Effects of Altered Avatars on Perception-Action in Virtual Reality

Brian Day
Clemson University, bmday15@gmail.com

EXAMINING THE EFFECTS OF ALTERED AVATARS ON PERCEPTION-ACTION IN VIRTUAL REALITY

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Psychology - Human Factors

---

by
Brian Day
August 2017

Accepted by:
Dr. Chris Pagano, Committee Chair
Dr. Rick Tyrrell
Dr. DeWayne Moore
Dr. Sabarish Babu

ABSTRACT

In virtual reality avatars are animated graphical representation of a person embedded in a virtual environment. Previous research has illustrated the benefits of having an avatar when perceiving aspects of virtual reality. We studied the effect that a non-faithful, or altered, avatar had on the perception of one's action capabilities in VR. In Experiment 1, one group of participants acted with a normal, or faithful, avatar and the other group of participants used an avatar with an extended arm, all in virtual reality. In Experiment 2, the same methodology and procedure was used as in Experiment 1, except only the calibration phase occurred in VR, while the remaining reaches were completed in the real world. All participants performed reaches to various distances. The results of these studies show that calibration to altered dimensions of avatars is possible after receiving feedback while acting with the altered avatar. Further, calibration occurred more quickly when feedback was initially used to transition from a normal avatar to an altered avatar than when later transitioning from the altered avatar arm back to the normal avatar arm without feedback. The implications of these findings for training in virtual reality simulations and transfer back to the real world are also discussed.

DEDICATION

This dissertation is dedicated to my family, Michael, Ingrid, and Kendall Day, all of whom have provided unwavering support, encouragement, and love throughout my entire life. Thank you for everything – I love you guys so much. I would like to thank my committee chair, Dr. Chris Pagano, for his assistance and guidance throughout this project, from developing a topic and writing the proposal to collecting the data and writing the defense. I am sincerely grateful for everything that you have done for me during my time at Clemson. I also would like to thank my committee members for their support. I truly appreciate all of your assistance and guidance. Specifically, I would like to thank Dr. Moore for his time, encouragement, and mentorship regarding statistics. I would also like to thank, Leah Hartman, my lab mate, for her assistance in everything I have done at Clemson – we make quite the team. In addition, I want to thank my mentor, Dr. Harry Heft. Without your help, I would not have been able to find my passion for this subject matter and your guidance in my academic career is invaluable. Finally, I would once again like to thank my family and friends for their support and encouragement throughout the years I have been working on this project.

TABLE OF CONTENTS

Table of Contents (Continued)

LIST OF TABLES

List of Tables (Continued)

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

The concept of a body schema has existed in the literature for over 100 years (Head, 1920; Head and Holmes, 1911). A body schema is the representation of the body and its potential for action. It is typically believed that the body schema is learned early in life and is based on information provided by the proprioceptive, vestibular, and kinesthetic senses (Iodice, Scuderi, Saggini, & Pezzulo, 2015). Originally, Head and Holmes (1911) postulated that any changes to the body and its action capabilities are compared to a body schema stored in memory. More recently, it has been hypothesized that the body schema is neither innate nor learned. Rather, the body schema is perceived. Accepting the hypothesis that the body schema is fluid and malleable allows for a body schema that is continuously perceived as the body moves and is equipped with items (clothing, hand-held tools, etc.) (Pagano & Turvey, 1998). A body schema is malleable in that it can be adjusted due to permanent or temporary changes made to the body or the body's abilities over the course of the lifespan. Over short time scales, people equip themselves with tools which requires calibration to new action capabilities. Over long time scales, the body grows and develops, which requires calibration as well. Iodice et al. (2015) found that while changes to bodily dimensions can result in adopting a new body schema; this is a relatively long and slow process. The process of calibration to the new capabilities, however, occurs much more quickly. Based on this finding, it seems that the malleability of the body schema is not durable when changes to the body are not permanent or cemented in to the perception-action system of the actor.

Humans frequently extend or augment their action capabilities through tool usage, which can be regarded as short-term changes to the body. Previous research has supported the idea that objects attached to the body, such as tools, are perceived as functional extensions of the body (Wagman & Chemero, 2014). The extension of the body through tool usage aides actors in both perceptual and behavioral tasks. The phenomenon of perceiving aspects of a distal surface by means of a handheld tool is also known as extended haptic perception (Burton, 1993; Carello, Fitzpatrick, and Turvey, 1992). Some investigators have proposed that through projecting sensations out to the distal end of a hand-held tool and associating the sensations with movements of the body, the mind is able to build a mental representation of the spatial layout of the body's current configuration (see Berti and Frassinetti, 2000; Cardinali et al., 2009; Lotze, 1856/1885, 1885/1973; Pagano and Turvey, 1993). Maravita and Iriki (2004) confirmed that use of a tool that functionally increases reaching ability causes an extension of the space that is perceived as reachable. Interestingly, the extension of perceived space persisted after the actor had discontinued using the tool to reach. It can be suggested that the functional increasing in reaching distance was incorporated into the body schema.

Other investigators, such as Gibson (1966) and Merleau-Ponty (1962), have hypothesized that actors can perceive environmental properties by means of non-innervated appendages (i.e. tools) because they do not simply perceive the tool, they attune to information specific to what is at the end of the tool. Attachments to the body are experienced just as parts of the body are (Pagano & Turvey, 1995, 1998), and numerous studies have supported these claims (Bongers, Michaels, & Smitsman, 2004;

Fitzpatrick, Carello, & Turvey, 1994; Wagman, Caputo, & Stoffregen, 2016; Wagman & Taylor 2005; Witt, Proffitt, Epstein, 2005).

As an example, it is frustrating to fail to reach an object that is just out of reach on a high shelf. Fortunately, there are many things we can do to obtain the object in question. For instance, we can stand on a step stool, use a grabber that extends our reach, or have a taller person reach the object for us. This highlights the idea that successful action requires that we be in tune with our action capabilities. For successful action one must be able to perceive what is or is not possible in the environment (Lessard, Linkenauger, & Proffitt, 2009). Luckily, human beings are quite good at perceiving what we can or cannot do in the environment, meaning we are quite good at perceiving affordances (Gibson, 1979).

According to James Gibson, "The perceiving of an affordance is…a process of perceiving a value-rich ecological object. Any substance, any surface, any layout has some affordance for benefit or injury to someone. Physics may be value-free, but ecology is not" (Gibson, 1979, p. 140). Affordances are the inherently meaningful aspects of the organism-environment system. Affordances are the relations between features of the environment and abilities of a person that make particular activities possible (Chemero, 2003; Gibson, 1979; Turvey, 1992). While affordances do not exist solely within the organism, they simultaneously exist both as a relational property in the dynamic and reciprocal organism-environment relationship *and* in the environment (see Heft, 2017; Gibson, 1966).

For example, if an object is too large in comparison to a person's hand size, then grasping the object is not afforded. In this case, the person would likely need to use two hands to pick up the object. Similarly, one can work comfortably at a desk if the heights of the surfaces (seat pan height, desk level height, location of objects on the desk, etc.) are correct relative to the physical dimensions and capabilities of the body. However, affordances are different for different people. An object that is within reach for a fully-grown adult may not afford reaching-to for a child, or someone with short arms. Thus, affordances are not situated in the environment (considered as separate from an individual) and they are not situated in the individual (considered as separate from the environment). Rather, they are relational properties.

Findings from the last thirty years of laboratory investigations in ecological psychology have established that all types of organisms are able to perceive their surroundings in terms of the opportunities for action that are afforded (Heft, 1993, Mark, 1987; Wagman, Thomas, McBride, & Day, 2013, Warren & Whang, 1987). However, the relationship between capabilities of an actor and environmental features continually changes over short and long time scales. Over the course of seconds, objects in the environment tend to move, which changes the possible actions for an actor. Similarly, people can become fatigued after acting for extended periods of time, which causes a change in action capabilities, just as when the addition or subtraction of carried loads can alter one's action capabilities. Over longer time scales, as the human body grows and develops these changes are coupled with changes in strength and coordination. Practice (or the lack thereof) can also alter action capabilities.

This ability of our perception-action system is no more apparent than at times when our capabilities change. Due to the fact that the affordances for a person in a given environment are constantly changing, affordances are dynamic (Fajen, Riley, & Turvey, 2009; Wagman, Higuchi, & Taheny, 2014). Fortunately, our perception-action system is flexible enough to adapt to changes in action capabilities. According to Welch (1986) adaptation is the semi-permanent perceptual-motor change that minimizes or eliminates a discrepancy between sensory modalities, within a sensory modality, or the errors in behavior due to the discrepancy. Traditionally, adaptation has been studied using prism goggles to induce a discrepancy between the visual information and the proprioceptive information specifying the location of the arm or a target in space (e.g., Welch, 1986). Before exposure to prismatic goggles, participants have no trouble locating their arm or pointing directly at a target in their environment. Then, once participants are exposed to the effects of the prism goggles a discrepancy between their visual and proprioceptive sense is created. Due to this discrepancy, participants are unable to point or reach directly at targets in their environment. But over a series of trials in the exposure phase, participants are able to correct their actions to become accurate. Finally, once the prism goggles are removed, participants are again unable to reach directly at targets in the environment but can slowly correct their actions. Overall, prism adaptation was traditionally measured using after-effects, while analyses of the exposure phase were of less importance. Interestingly, it usually takes fewer trials to recalibrate in the final phase than it does to calibrate to the prism goggles in the exposure phase. Generally, research into the process of adaptation has highlighted the plasticity of the human perception-action system in responding to discrepancies.

However, adaptation occurs relatively slowly (Bingham & Romack, 1999). Other investigations have demonstrated the ability of human actors to rapidly calibrate to discrepancies much quicker than the process of adaptation (Bingham & Romack, 1999; Welch, Bridgeman, Anand, & Browman, 1993). Specifically, Bingham and Romack (1999) hypothesized that active reaching allows for the discrimination of new sources of visual information and for the actor to calibrate to the appropriate information. Once the new sources of information become salient, the perception-action system can rapidly recalibrate to the altered action capabilities, and their results support this hypothesis. Bingham and Romack (1999) argue that the ability to rapidly adjust is more functionally adaptive than the traditional process of slower adaptation. Thus, a crucial process for active organisms to engage in is calibration, which is the process by which the perception of affordances and the execution of actions become scaled to the (changing) relationship between environmental features and action capabilities (Bingham & Pagano, 1998; Fajen, 2005, Withagen & Michaels, 2004; 2007).

In previous literature, the terms adaptation and calibration have at times been used interchangeably (Mon-Williams & Bingham, 2007) and other times differentiated, as discussed above. Sometimes, the term adaptation has been referred to as the process by which actors adjust to manipulations of embodied sensory units (Bingham & Pagano, 1998; Bingham, Pan, & Mon-Williams, 2014; Coats, Pan, & Bingham, 2014), while calibration has been referred to as the process by which actors adjust to changes to action units. Overall, the meaning of both terms is not generally agreed upon in the literature (van Andel, Cole, & Pepping, 2017). Perhaps adaptation is more general than calibration, in that the proper scaling of perception and action (i.e. calibration) is just one form of adaptation. In this sense, calibration is a special form of or case of the more general process of adaption. Calibration, just like adaptation, involves aligning perception and/or action to the proper scale so as to be accurate and functional for the actor relative to some standard that is sensed simultaneously. Terms like 'tuning' or attunement are distinct from adaptation, because attunement refers to the selection of variable(s) within the ambient energy array to use as the systems input for the particular task at hand.

From an ecological standpoint, there is no use in differentiating between adaptation and calibration. Nothing is gained by splitting adaptation from calibration, as both refer to an adjustment between two measures. According to the ecological approach, organisms perceive to inform action, and act to inform perception. Referring to separate processes like sensory adaptation (i.e. the input) and action calibration (i.e. the output) is a harmful reduction when attempting to understand behavior at the level of scale where organisms act. Both processes relate to perception-action learning, and for the purposes of this paper this process will be referred to as calibration only.

One of the earliest studies to investigate calibration was carried out by Mark (1987). In the initial part of the study, participants were presented with a chair whose seat height could be adjusted. Participants were asked to judge whether the presented chair was low enough for them to sit on in a given trial. The results of the initial phase of the study indicated that people give very accurate estimates of their ability to sit on a chair; meaning when the ratio of seat height to leg length was below a certain point participants always reported the chair as being "sit-on-able." However, when the ratio of seat height to leg length was above a certain critical point participants began to judge the chair as not sit-on-able. Interestingly, Mark then manipulated participants sitting ability, by attaching 10-cm tall wooden blocks to the soles of their feet. Now, participants could sit on chairs that were 10 cm taller than before. At no point during this phase of the study were participants allowed to look at the blocks attached to the feet. Further, participants were not allowed to practice sitting either. Again, participants were asked to judge if a chair was sit-on-able. The major finding of the second phase of the study indicated that participants quickly adjusted, or calibrated, to their new capabilities by altering their judgments of what was sit-on-able. Afterwards, Mark asked participants to judge how tall the block attachments were, and these estimates tended to be inaccurate. The findings suggest that the human perception-action system has the ability to calibrate to altered capabilities without knowledge of the specific alterations.

Other research has shown that the visual perception of affordances for a particular behavior can become calibrated after explicit practice performing that behavior (Franchak, van der Zalm, & Adolph, 2010; Wagman, 2012), or even a related behavior (Stoffregen, Yang, Giveans, Flanagan, & Bardy, 2009; Wagman et al., 2014). Most interestingly, other empirical investigations have revealed that calibration for a certain behavior can also occur following simply perceiving affordances for that behavior. This means that calibration can occur in the absence of physical activity, explicit feedback, or knowledge of results (Mark, 1987; Mark, Balliett, Craver, Douglas, and Fox, 1990; Ramenzoni, Davis, Riley, & Shockley, 2010).

Recent research has investigated whether calibration occurs in the same manner in virtual reality. Altenhoff et al. (2012) studied the effect of visual and haptic feedback on depth estimations in VR. Participants who received visual and haptic feedback made more accurate distance estimates after the calibration phase, suggesting that calibration of depth estimates can occur in VR. Ebrahimi, Altenhoff, Pagano, and Babu (2014) showed that participants calibrated to perturbed visual distances, meaning that if their visually presented end effector was shown to be nearer, participants believed they were underestimating their reaches to targets and after feedback began to overestimate. Similarly, Ebrahimi et al. (2015) found that depth judgments in VR are more accurate when scaled to visual and haptic feedback during closed-loop reaches than depth judgments made in an open-loop manner in the real world. This finding is important because it suggests that visual feedback is necessary for the calibration of actions in VR, and that congruent visuo-haptic feedback is most effective for calibration. Building off of this finding, Ebrahimi, Babu, Pagano, and Jorg (2016) revealed that the presence of accurate visual feedback alone is sufficient for calibration of reaching actions to occur in VR. Interestingly, none of these studies incorporated a fully rendered avatar for the user in VR.

With current technology, virtual environments allow people to perform tasks and actions that may not be possible or feasible in the real world, because of safety concerns or limited resources. Virtual reality systems often represent the user as an avatar.

Avatars are animated graphical representations of people embedded in virtual environments (Lin, Reiser, & Bodenheimer, 2015). Until recently, most studies utilizing immersive virtual environments (IVE) did not include human avatars due to the technological difficulty in rendering a realistic avatar (McManus et al., 2011). But this trend has begun to change.

Previous work has highlighted the importance of sense of presence, specifically a sense of embodiment, when acting with a self-avatar in VR (Kilteni, Groten, & Slater, 2012). A sense of embodiment towards a particular body, such as a self-avatar, is the sense that emerges when the properties of the new body are perceived as the properties of one's own biological body (Kilteni et al., 2012). For example, objects such as virtual limbs can be experienced as part of one's own body when specific types of synchronous multi-sensory and sensorimotor simulation exist (Slater, Perez-Marcos, Ehrsson, & Sanchez-Vives, 2008; Slater, Perez-Marcos, Ehrsson, & Sanchez-Vives, 2009). Through the presence of synchronous visual-tactile stimulation, Slater et al. (2008) showed that the rubber hand illusion could be replicated in virtual reality. Their results indicate that the visual-tactile synchrony is important in VR, especially for those which require some kind of interaction with the virtual environment. Embodiment can also be induced through first-person viewpoint of the virtual body where there is a visuo-motor synchrony between the real body and virtual representation (Bankour, Domna, Groten, & Slater, 2013; Maselli & Slater, 2013). Overall, exposure to multisensory and sensorimotor information while acting with an avatar can result in a sense of embodiment even when the avatar has body dimensions different than our own (Kilteni, Normand, Sanchez-Vives, & Slater, 2012; Kokkinara, Slater, Lopez-Moliner, 2015).

Previous studies have investigated the effect that an altered avatar has on an actor's perception. The results of Slater et al. (2008) support the idea that activities that are impossible to do in the real world, such as altering the length of limbs or size of one's body in real time, can be done in VR, as simulated objects can be incorporated into the body representation and treated as part of the participant. In some cases, an avatar may faithfully represent the anthropometric dimensions of the user (Lin, Rieser, & Bodenheimer, 2012; McManus et al., 2011; Mohler, Creem-Regehr, Thompson, and Bülthoff, 2010). In other cases, the avatar may not always be a direct reproduction of the user – this can occur mistakenly, or the avatar may be purposively different (Jun, Stefanucci, Creem-Regehr, Geuss, & Thompson, 2015; Leyrer, Linkenauger, Bülthoff, Kloos, & Mohler, 2011).

Leyrer et al. (2011) manipulated the eye height of the avatar participants used while asking them to judge distances in virtual reality. As part of their study, some participants had the eye height associated with their avatar increased by 50 cm, some participants did not have their eye height manipulated, while others had their eye height decreased by 50 cm. They found that participants with an avatar whose eye height was increased by 50 cm perceived distances as shorter in comparison to participants who viewed distances through a shorter eye height. Interestingly, participants who had their eye height decreased by 50 centimeters did not show an increase in distance perception, nor was there a difference in distance perception between the shortened eye height group and the normal eye height group. From the standpoint of calibration, this asymmetrical finding is quite intriguing, as it suggests that calibration to altered dimensions of an avatar is not automatic. Perhaps most important, is the fact that changes in presented eye height only affected judgments of distance when ownership of the avatar was felt, suggesting that participants who did not feel as if the avatar was their own were unwilling to calibrate to the altered dimensions.

The above-mentioned experiment further emphasized the importance of providing an avatar for the accurate perception of virtual environments (Creem-Regehr, Stefanucci, & Thompson, 2015). Not only can the virtual environment be perceived as smaller, but the perception of action capabilities in the virtual environment can be altered as well, based on the presence or absence of an avatar (Mohler et al., 2010; Renner, Velichkovsky, & Helmert, 2013; Ries, Interrante, Kaeding, & Phillips, 2009). Lin, Reiser, and Bodenheimer (2013) investigated whether the presence of an avatar has an effect on perception of action capabilities. In their experiment, participants were tasked with judging their ability to safely step off of a ledge without falling. Half of the participants were presented with an avatar that faithfully represented their own body dimensions, while the other half of participants were not provided with an avatar. Participants with an avatar estimated that they could step off of ledges that were approximately 25% of their eye height, while participants with no avatar estimated they could step off of heights up to 50% of their eye height (from which is too tall to step safely). Further, Lin, Reiser and Bodenheimer (2015) found that providing a self-avatar in a virtual environment generates action judgments that are not significantly different from action judgments made in the real world. The authors conclude that having participants perceive their action capabilities in a virtual environment with the presence of an avatar allows for a fairly accurate judgment of the fidelity of that virtual environment.

Overall, it seems that by manipulating an avatar, perception of the virtual environment can be altered, suggesting that calibration can occur in virtual reality. Specifically, by increasing the eye height of an avatar, the virtual environment is effectively condensed (Creem-Regehr, Stefanucci, & Thompson, 2015). Additionally, the results of Lin et al. (2013; 2015) suggest that providing faithful avatars allows for people to perceive critical information when deciding how to act in virtual environments.

Jun et al. (2015) investigated perception of action capabilities in VR with avatars that were altered in some way. Their participants were tasked with judging the width of a gap in virtual reality and were also asked to judge if they could safely cross that gap. All participants were shown only the disembodied feet of an avatar. Participants were randomly assigned to either the small or large foot condition. In the small foot condition, the presented feet were only 50% of the standard American male foot, and in the large foot condition the presented feet were 200% of the standard foot. The researchers found that when participants viewed the environment with large feet they judged distances as being relatively shorter and indicated that they could step over relatively wider gaps in comparison to the small feet condition. In addition to showing an effect on judged distances in virtual reality, the authors were able to show that perceptions of action capabilities could be altered due to manipulations in the size of portions of the presented avatar (Jun et al., 2015).

Another study, performed by Linkenauger, Leyrer, Bülthoff, & Mohler (2013) manipulated the size of a virtual hand to observe the effect on participants' ability to judge the size of an object placed next to the hand. The size of the hand was either small, medium, or large and connected to an arm that did not change in size. When virtual hand size was small, participants judged the object as being larger. Conversely, when virtual hand size was large, participants judged the object to be smaller. In addition to giving judgments of the size of an object, participants were asked to report if they could grasp objects of different sizes. When presented with different sized hands in VR, the pattern of results regarding grasp-ability of objects was the exact same as the pattern of results regarding object size.

Experiment One

One example of a task that requires near-constant calibration is determining what is within reach of the body, as humans continually interact with hand held tools and graspable objects of various dimensions. Once a tool or object is grasped, the capabilities of an actor change. Specifically, what is now within reachable space changes. For example, when using a tool, action possibilities in the environment generally increase and humans calibrate to this change in capabilities quite readily (Witt and Proffitt, 2008). Ebrahimi et al. (2016) urged future researchers to investigate the impact of an immersive self-avatar in IVE, and examine its effects on human reach actions.

Previous research has not specifically addressed the question of whether extending the dimensions of the arm of an avatar is effectively the same as using a tool that increases reaching ability in the real world. Nor has previous research investigated the effect of having participants calibrate to altered action capabilities in the middle of an experiment. Specifically, the current study will address the following questions: If an avatar possesses different anthropometric dimensions than your body, can your perception-action system quickly calibrate (i.e., after a limited number of trials) to the dimensions of the altered avatar when attempting a simple action? If the calibration is not quick, does exposure to using the altered avatar in VR facilitate calibration? If participants are able to calibrate to the altered dimensions and action capabilities of the lengthened avatar, this will have significant ramifications for understanding the malleability of the body schema in VR.

The present research has ramifications outside of virtual environments as well. For instance, the present work is similar to the idea of accepting a limb that is bigger (or smaller) than your own limb, such as when amputees receive artificial limbs. These artificial limbs may or may not be the exact same size as their lost limb. Further, the present work has ramifications for robot teleoperation, in that if humans can readily use avatar limbs that are longer than their own, then the same phenomena might hold true for robotic teleoperation.

The research questions were investigated through the first experiment. The first experiment contained two primary avatar types (altered avatar vs. normal avatar) and utilized three blocks of experimental trials. The tasks that constituted each block involved participants performing reaches to virtual targets at various distances.

Hypotheses

The current study has four primary hypotheses. Based on previous findings, in terms of absolute error as the primary dependent variable (ABS(estimated distance – target distance)), we predicted that calibration to an altered avatar would occur but it would not be instantaneous. Rather, we expected a linear improvement to a critical point in error over trials in the calibration phase. Next, based on the findings in the adaptation and calibration literature that have revealed malleability of the body schema over periods of brief exposure to altered action capabilities, we predicted that calibration to the altered avatar with feedback would occur more quickly than reversion from the altered avatar back to a faithful avatar arm length without feedback. This would be evidenced by steeper linear improvement in error across trials with the altered avatar in the calibration phase than in the posttest. But reversion back to the user's normal body schema would still occur in the posttest. Reversion will be defined as exhibiting absolute error that matches the error demonstrated by the normal avatar group. However, reversion would occur less quickly in the posttest than calibration occurred in the pretest, resulting in a less steep slope predicting accuracy across trials than in the calibration phase. Similarly, we predicted that in the posttest, participants in the altered avatar condition would exhibit greater under-reaches, and would reach to more unreachable targets than participants in the normal avatar condition until after fifty trials have occurred.

CHAPTER TWO

METHOD

Participants

Just as in power analyses for traditional statistical techniques, estimating power in a multilevel study still deals with investigating the power of a statistical test as a function of Type I error rate, sample size, and effect size. Two other considerations for estimating power in a multilevel study are the sample size of Level 2 units compared to the sample size of Level 1 units, and the intraclass correlation (ICC). Power estimation in MLM is a complex procedure because it requires additional assumptions due to nesting and the Level 1 and Level 2 estimates. Simulations have been run manipulating the n at Level 1 and N at Level 2 to determine the standard error in various scenarios.

Using absolute error as the dependent variable of interest, based off previous research (Day, Ebrahimi, Hartman, Pagano, & Babu, under review) our estimated ICC was 0.15. To be conservative we followed guidelines presented in Hox, Moerbeek, and van de Schoot (2010) by estimating the design effect based on both 20 and 24 participants. Using these participant estimates and 130 L1 units, the design effect was 20.35. The effective sample size was 128 for 20 participants, and 153 for 24 participants. For Cohen's medium effect size (f^2) of .15, with an alpha level of 0.05 and seven IVs, power would be between 0.88 and 0.94, for 20 and 24 participants respectively.

In order to have power of at least 0.80, we would need an effective sample size of at least 105. We chose an effective sample size that falls comfortably between these two estimated effective sample sizes. Assuming the ICC of our dependent variable does not exceed 0.25 (which is likely), by recruiting 26 participants (an N at L2 of 26), each of which will complete 130 trials (an n at L1 of 130), we would far exceed an estimation of power at 0.80.

Twenty-eight undergraduate students (22 females and 6 males, *M* = 18.68, *SD* = 0.72) from Clemson University participated in this experiment. Data from two participants were discarded due to a malfunction of the tracking system. Participants were required to be right handed as all equipment used was for right-handed participants. All participants received credit in their psychology courses in exchange for participation. As participants entered the testing area, they were given a brief overview of the purpose of the experiment and informed consent was obtained. Participants with a history of stroke or epilepsy were ineligible to participate in this experiment. If participants needed glasses or corrective lenses they were asked to wear those while participating. Participants were administered tests for visual pathologies (such as refractive error or stereo blindness) before completing any trials. If participants failed these tests they were unable to participate in the experiment. Participants were randomly assigned to either the altered avatar condition or normal avatar condition. A between-subject approach was used for the primary manipulation to allow for direct comparisons between participants acting with avatars with different action capabilities. A between-subject approach was favored so as to avoid having participants attend multiple experimental sessions over various days, and to avoid practice effect associated with within-subject designs (Jun et al., 2015; Lin et al., 2013; Linkenauger et al., 2013).

Design

The current experiment utilized a 2 (Avatar Type: Altered avatar vs. Normal avatar) by 3 (Phase: PreTest, Calibration, PostTest) mixed groups design. Avatar type was a between-subject variable and phase was a within-subject variable. The normal avatar condition involved use of an avatar's arm that was directly proportional to the dimensions of the user's own arm. The altered avatar condition involved use of an avatar whose arm length, and thus reaching capabilities, were increased by 30 cm.

Materials and Apparatus

Figure 1 depicts the apparatus that was used to present the VR. Participants were seated in a wooden chair, which was situated approximately 20 cm from the edge of the wooden table. The tabletop was 50 cm wide by 130 cm long, and was 76.2 cm tall (which is standard table height). The center of the table was aligned with the midpoint between the participants' right eye and right shoulder. Participants were outfitted with five Pohlemus sensors on the forehead, neck, right shoulder, right elbow, and on the hand-held tool. Aside from the sensor on the forehead and on the tool, the other three sensors were all placed on the bony protrusions at those points on the body. The base for the Pohlemus system was located underneath the table and out of view of the participants. The virtual environment, which was a recreation of the same room, was displayed using a HTC VIVE head mounted display (HMD), which is a binocular display system that displays stereo information by presenting different information to each eye, with a combined resolution of 2160 x 1200 pixels, a 90 Hz refresh rate, and a 110-degree field of view. A virtual table and chair, whose dimensions and positions were the same as the real table and chair, were placed centrally in the virtual room. See Figures 1 and 2.

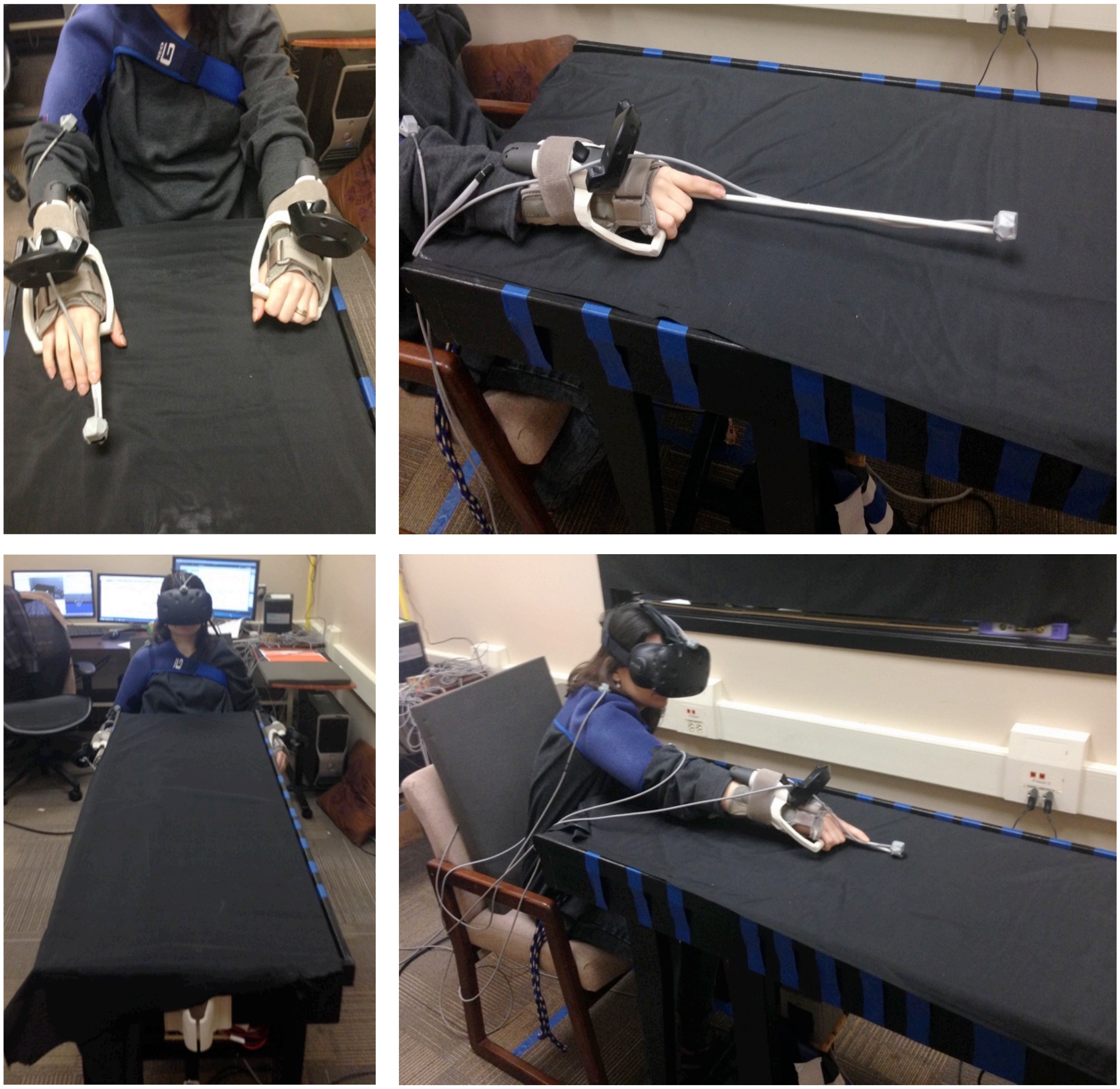

Figure 1. Top: A view of the VIVE controllers, wrist worn mount, and both tools (normal and long). Bottom: The table apparatus. This configuration was also rendered in virtual reality. Participants were asked to reach to targets presented at the horizontal midpoint of the table.

As mentioned previously, two types of avatars were used (see Figure 2). In any given block of trials participants were asked to reach for a target with their right arm and hand. Participants were asked to reach as quickly and as accurately as possible. As soon as participants initiated their reach, their field of view was grayed out, so that each reach was completed in an open-loop manner. The normal avatar did not increase the reaching capabilities of the actor whereas the altered avatar increased reaching capabilities by 30 cm. In the real world, participants were given a Vive controller to hold. See Figure 1.

The Vive controller was 26.5 cm long from base to tip, 3 cm wide at the base of the handle, 5 cm wide at the top of the handle, 3 cm deep at the handle, and is 12 cm wide at its widest point. The Vive controller was mounted on a plastic mold affixed to the top of both of a participant's wrists, as seen in Figure 1. The wrist brace allowed for the wrist to remain in a consistent orientation across all trials and across all participants. Mounting the VIVE controller on top of the wrist brace allowed the experimenters to accurately model participants wrist position and hand position in VR. In this way, the participants were presented with an avatar that accurately represented the orientation of their arms in the real world. Participants were unable to see their shoulders or upper arm segment while reaching in VR. The plastic mold designed to hold the controller also held a plastic rod with a rubber tip. When participants were reaching with the normal avatar in VR, a 10-cm plastic rod was inserted into the mold. When participants were reaching with the altered avatar in VR, a 30-cm plastic rod was inserted into the mold.

Participant's head and hand movements in the real world were tracked and this information was used to update the image displayed in the HMD so that the head and hand movements of the avatar were consonant with participants' movements in the real world. Inverse kinematic (IK) algorithms were used to update the position of the forearm and upper arm segments based on the position of the head, shoulder, and hand.

Generally, IK can accurately predict the position of the arm segments, yet the algorithms are not always perfect so there was a chance for error in the positioning of the virtual arm (however, this did not commonly occur in the study). By outfitting participants with the wrist brace, the orientation of their wrist and fingers was consistent for the entire experiment, and this position was maintained in the appearance of the virtual avatars wrist and fingers as well.

In this study, we substituted the long tool with the extension of the forearm of the participants. In the normal avatar condition, participants observed the self-avatar holding the short tool. In the altered avatar condition the length of the tool will be added to the forearm of the avatar. For participants in the altered avatar condition, the IK algorithms elongated the upper arm and forearm segments by a cumulative 30 cm. No other dimensions of the arm or hand were altered.

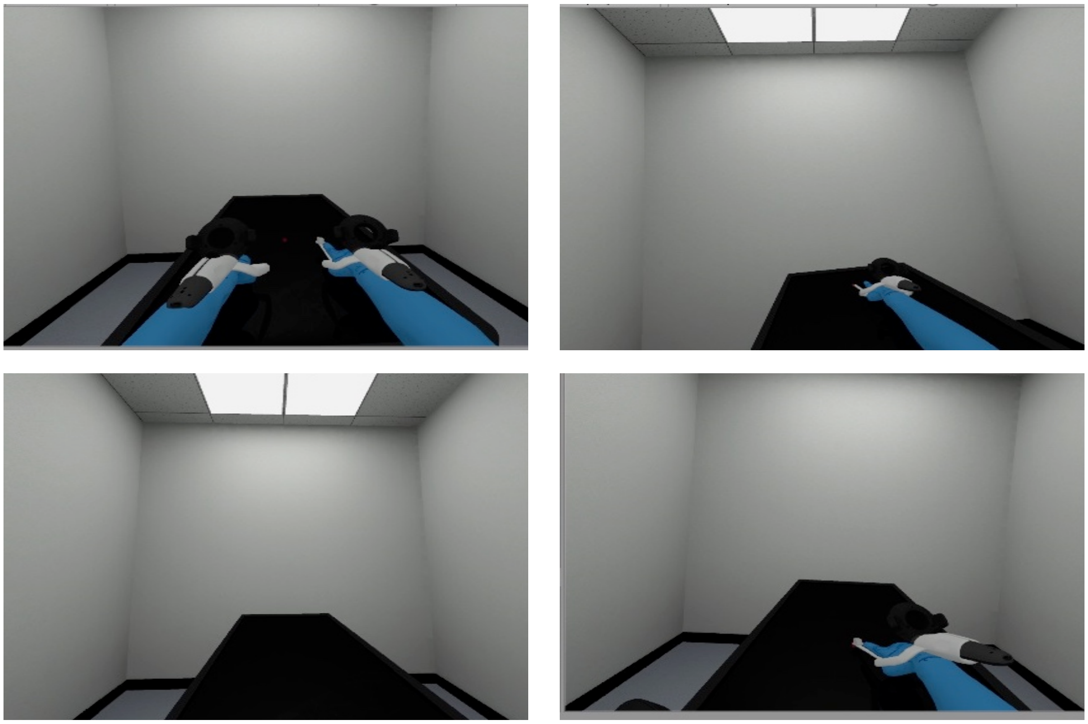

Figure 2. The rendering of the virtual environment and the avatar as seen by participants. Each picture corresponds to the virtual scene the participant would see for each of the four images in Figure 1, respectively. Starting on the top left and going clockwise - A) Both hands extended. B) Altered avatar reaching for a target. C) Resting position. D) Normal avatar reaching for a target.

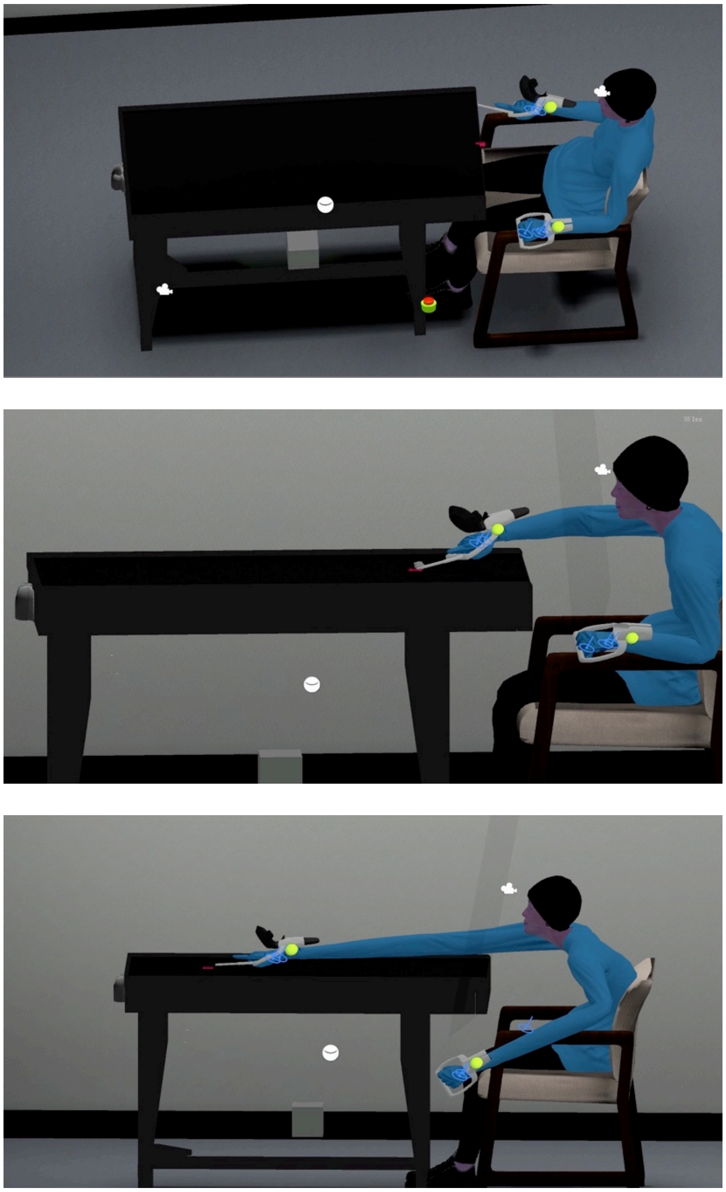

Figure 3. The rendering of the virtual environment and the avatar as seen by experimenter. Each picture corresponds to the virtual scene the experimenter would see for A) Resting position. B) Normal avatar reach. C) Altered avatar reach.

Participants were asked to reach for a visual target in virtual reality. For any trial, the target consisted of a virtual representation of three luminous LEDs. The middle LED corresponded to the target distance. With the other two LEDs luminous the length of the target area was 3 cm. In the pre- and posttest, targets were presented at 13 different distances, ranging from 20.5 cm to 121.5cm. The difference between each target was approximately eight cm. In the calibration phase, targets were presented at 9 new distances, ranging from 17.4 cm to 114 cm, and the distance between each target was approximately 10.5 cm. Every target was presented randomly, and each target distance was presented five times for a total of 65 reaches in the pre- and posttest. The target distances presented in the calibration phase were presented five times each as well.

Procedure

As participants entered the testing area, they were given a brief overview of the purpose of the experiment and informed consent was obtained. Participants were administered the Stereo Fly Test (Stereo Optical, Chicago, IL), which tested gross stereopsis and fine depth perception. Participants were then administered a test to determine inter-pupillary distance (IPD) to help ensure the VIVE VR headset was properly adjusted to each participant. As detailed by Willemsen et al. (2008), the IPD test called for participants to look into a mirror from a set distance and mark the location of each pupil in the mirror. The experimenter then measured the distance between the two marks. The measured IPD was used to set the inter-ocular distance on the VR headset accordingly. By ensuring that the IPD of the VR headset was adjusted correctly for each participant, retinal disparity and vergence would remain intact when participants were viewing the virtual environment.

All participants were asked to sit on the wooden chair at one end of the wooden table. Various motion sensors were placed on the participant through the use of a long sleeve shirt. The sensors were attached to the shirt with Velcro, and the cords for the sensors were strapped to the arm of the participant. The straps helped keep the shirt tight to the arm of the participant so as to not interfere with their reach, and the straps helped to keep the wires of the tracking system from pulling on the system. The physical location of each sensor was measured before and after data collection to ensure that the sensors did not move over the course of the experiment.

Before putting on the HMD, the experimenter demonstrated the types of reaches that were appropriate in the experiment. Then, after putting on the headset, but before any trials occurred, the participant engaged in three tasks to familiarize themselves with being in VR. Participants were able to see their self-avatar in a mirror in the VR simulation. The purpose of completing these tasks was also to induce a feeling of body ownership with the self-avatar. The tasks were based upon those frequently used by Slater in his research on presence in VR (Bankour et al., 2013; Kilteni, Groten, & Slater, 2012; Maselli & Slater, 2013). The first task required participants to bring their arms up to their side and move them around so they could see how the movements of their body caused the avatar to move simultaneously. The second task had participants stretch their arms out straight in front of them and rotate their wrists. Lastly, participants were asked to stretch their arms up over their head and move their arms around.

Participants were instructed to reach as quickly and as accurately as possible on each trial. The major restriction participants had was they needed to remain seated (i.e., keep their weight on the seat pan) and keep their feet flat on the floor during each reach. During the course of the actual reach participants could engage their arm only, or they could engage their entire upper body (i.e. bending at the waist to reach farther).

Regardless of phase, each trial began with the participant resting their right forearm on the armrest of the chair and their back against the back of the chair. Participants were instructed that this was the starting point for each trial. To ensure uniformity in starting positions across participants, it was emphasized to participants that this starting posture is critical for the study. Across all phases, participants were instructed to reach out as quickly and as accurately as possible, and place the tip of the stylus as close to the center of the target at possible.

Pretest

In the pretest, participants were instructed to reach to the target that appeared on the table at various distances from them. As part of each trial the participant was asked to make a judgment if they could reach the target or not. If the participants answered in the affirmative (by saying "yes"), they were then instructed to initiate a reach. To ensure that participants could not see the target while reaching or receive informative feedback about their reach, at the initiation of their reach, participants were shown a grey screen to simulate closing their eyes. After attempting to reach the target, participants were instructed to return their hand and arm to the starting point to begin the next trial. If participants did not believe they could reach the target they were instructed to say "no", and the next target distance was presented. Regardless of condition, all participants performed the pretest with a normal avatar (i.e., accurately customized for their arm length). In this phase, participants only received haptic feedback when the controller they are wielding in the real world contacts the surface of the table, but this feedback did not inform them about how close their reach was to the target.

Thirteen different target distances were presented to each participant in the pretest. The distances ranged from approximately 20 cm to 120 cm away from the participant. Each distance was separated by approximately 8 cm. Targets were presented randomly, and each target distance was presented five times, for a total of 65 trials.

Calibration phase

After the pretest, participants completed the calibration phase. In the calibration phase, participants performed fewer reaches to fewer targets than in either the pretest or posttest phases. Nine new distances that had not been presented in the pretest (and were not presented in the posttest) were presented in the calibration phase. Each of the nine targets was presented five times for a total of 45 trials, and all targets were presented randomly.

The task in the calibration phase was very similar to the pretest, except in this phase participants could see the result of their reaches. After being shown a target, participants still gave a judgment if they could reach to the target or not. Regardless of their response, participants were asked to reach to the target when the screen went blank. Once the initial reach was made, the virtual scene was restored to the headset so participants could see the result of their reach. At this point, if the target was within reach, participants were asked to adjust their reach to the center of the target area and hold there for one second before returning their hand to the starting position. If the target was clearly out of reach participants returned to the starting position.

The primary manipulation of the experiment occurred in the calibration phase. Participants in the normal avatar condition continued reaching with a normal avatar.

However, participants in the altered avatar condition reached with an avatar arm that was 30 cm longer than their normal avatar. For participants in the altered avatar condition, a plastic rod that increased reach by 30 cm was substituted for the plastic rod that was used in the pretest. Participants were not told of this functional increase in reaching ability. In the calibration phase, participants received haptic feedback from when the (unseen) physical controller brace they were wielding in the real world contacted the surface of the (physical) table. As stated above, once contact was made with the table, participants were shown the virtual scene again and told to adjust their reach so the end of the virtually presented hand was in the center of the target, thus receiving visual feedback as well.

Posttest

The posttest was identical to the pretest. Importantly, the experimenters ensured there was minimal delay (i.e. no longer than 45 seconds for any participant) between the calibration phase and the posttest. By doing so, we hoped to preserve the just modified action capabilities of the avatar for the posttest, as a long delay between these two phases might cause the calibration to disappear.

Post Data Collection

After the conclusion of data collection, the experimenter again measured various aspects of the participant's arm to ensure that the positions of the sensors did not move over the course of the experiment. If a sensor was found to have moved more than 5 cm, the data for that participant was deemed unusable and not included in the statistical analysis. In addition, the participant was asked to perform three reaches with their arm only (i.e. reaching their arm straight out as far as they could without engaging their shoulder or back) and three maximum reaches with their entire upper body (i.e. reaching as far as they possible could and touching the table with no restrictions other than remaining seated in the chair with their feet flat on the floor). Participants were given a brief questionnaire designed to measure the degree of body ownership they felt over the avatar in VR (see the appendix). The questionnaire contained items similar to those used in previous research (see Maselli & Slater, 2013; Slater et al., 2009, Slater et al., 2010). A manipulation check was also administered to participants. They were asked if they noticed anything odd that occurred during the course of the experiment. Lastly, participants were asked about their previous use of VR simulations.

CHAPTER THREE

RESULTS

Body Ownership

Mean responses on each of the six items in the body ownership questionnaire were compared between groups. As can be seen in Table 1 there were no significant differences in feelings of body ownership between participants in the normal and altered avatar groups.

Table 1
Means, Standard Deviations, Standard Errors, and Significance Values for responses to the Body Ownership Questionnaire.

	Condition	N	Mean	Std. Deviation	Std. Error Mean	Significance
Question1	Normal avatar	13	7.0769	2.17798	.60406	.394
	Altered Avatar	15	6.1333	3.35659	.86667	
Question2	Normal avatar	13	7.3846	1.80455	.50049	.219
	Altered Avatar	15	6.1333	3.15926	.81572	
Question3	Normal avatar	13	8.4615	1.66410	.46154	.632
	Altered Avatar	15	8.1333	1.88478	.48665	
Question4	Normal avatar	13	3.7692	2.86222	.79384	.652
	Altered Avatar	15	3.2667	2.93906	.75886	
Question5	Normal avatar	13	3.0000	2.38048	.66023	.568
	Altered Avatar	15	3.4667	1.88478	.48665	
Question6	Normal avatar	13	7.3077	2.28709	.63432	.722
	Altered Avatar	15	6.9333	3.08143	.79562	

In response to the post-data collection manipulation check, 18 participants (64% of total participants) indicated that they noticed something odd during the course of the experiment while 10 participants (36% of total participants) indicated that nothing seemed odd. Of those 18 participants who responded that they noticed something odd during the experiment, only seven participants (39% of yes responders) mentioned something about the arm of the avatar being extended, manipulated, or larger. Common responses were, ‘color of the eyes and skin was off’, ‘it looked weird but not sure what’, or ‘arms were longer’. Broken down by avatar type, of the 13 participants in the normal avatar group, six participants (46%) indicated they noticed something odd and seven participants (54%) said they did not notice anything. Of the 15 participants in the altered avatar group, 12 participants (80%) indicated they noticed something odd and three participants (20%) said they did not notice anything. Of those 12, six participants (50%) specifically mentioned something about the arm of the avatar being extended, manipulated, or larger.

Transformation Variables

Categorical variables included condition (normal avatar group used as reference category), phase (pre-test used as reference category), and error direction (over-reach used as reference category). Figure 3 demonstrates the raw data in terms of reached distance (the distance to which participants reached with the tip of the tool) and presented target distance. The overall data is shown, as well as the data for each phase.

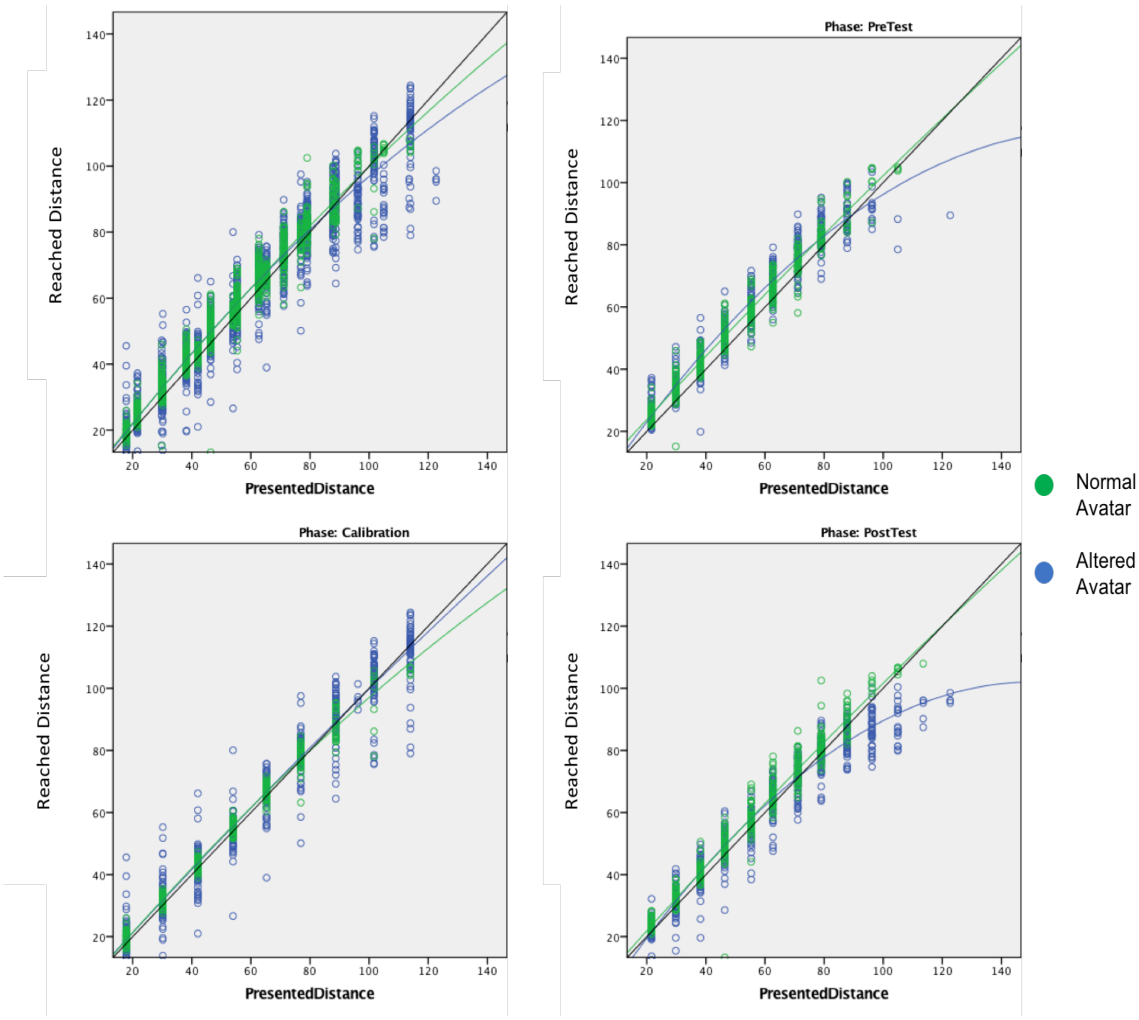

Figure 4. Estimated distance as a function of presented distance (clockwise from top left) a) overall b) pre-test c) calibration phase d) post-test. The solid black line in each graph represents perfect performance (y=1x+0).

Absolute error was calculated by taking the distance between the presented target distance and the estimated distance for each trial (error = reached distance – presented target distance), where negative values indicate under-reaching and positive values indicate over-reaching. Then, the absolute value of the error term was computed. Three binary variables were created using information contained in this error term. First, negative error values were coded as 0 (under-reach) and positive error values were coded as 1 (over-reach). Secondly, a correct judgment term was computed that evaluates whether participants correctly judged if the presented target distance was within their reach envelope. The correct judgment variable takes into account whether the target was within reach or not on a given trial and the participants' response on that trial. Attempting to reach to targets outside of the reach envelope or not reaching to targets that were within reach were coded as incorrect judgments (0). Reaching to distances that were within reach and not reaching to targets that were out of the reach envelope were coded as correct judgments (1). Lastly, regardless of correct judgment, if participants performed a reach, that trial was classified as 'action taken' (1). If they did not reach, that trial was classified as 'no action taken' (0). This variable is referred to as action taken. For example, if on a given trial the participant over-reached the target by reaching to a distance further than the target distance, that trial would be coded as 1 for over-reaching, 1 for correct judgment, and 1 for action taken. Conversely, if on a given trial the participant undershot the actual target distance by reaching too short to a target that was out of reach, this trial would be coded as 0 for under-reaching, 0 for correct judgment, and 1 for action taken.

Outlier Analysis

For each analysis, individual outlier analyses for full models were conducted. Residuals were obtained, standardized, and examined for any potential outliers that were outside of the normal distribution (Cohen et al., 2003). Outlier analysis was based on data visualization as well. Data points that were likely due to malfunctions in the tracking equipment and were not physically possible were removed for each specific analysis. In all of the analyses less than 2% of the trials were removed due to outliers.

Hierarchical Linear Modeling

The intraclass correlation (ICC) of the intercept only model (null model) was used to assess the overall nesting within participants for each of the main dependent variables (correct judgment and absolute error). Due to the repeated-measures design of the experiment, variables had significant nesting within participants. For example, the ICC of absolute error was approximately 15%. An ICC greater than 2-3% indicates nesting that demands a multilevel modeling approach (Bliese, 1998; Heck, Thomas, & Tabata, 2010). Multilevel modeling offers a more flexible approach to accurately modeling data produced in repeated-measures designs over traditional analyses such as a repeated-measures ANOVA (Cohen et al., 2003).

As previously stated, predictor variables for Level 1 were collected at each trial occasion (e.g. presented distance, presented distance quadratic, and action taken) and person level predictors (Level 2), were collected for variables such as condition. Interactions terms were also created which could be inter-level interactions (e.g., Level 1 by Level 1 *or* Level 2 by Level 2) or cross-level interactions (e.g., Level 1 by Level 2).

In multilevel modeling, effect sizes, also known as pseudo-R^2, are indexed by a measure of percent reduction in error variance. Level 1 error variance is indexed by a reduction in residual variance for Level 1 predictors. Level 2 error variance is indexed by a reduction in intercept variance for Level 2 predictors. Reduction in error variances for cross-level interactions (Level 1 by Level 2) is indexed by the percent reduction in the

Level 1 slope variance. The R^2 change is only calculated for significant effects, and the unique effects controlling for all other variables in the model.

Multilevel modeling relies on both general linear models and generalized linear models. Thus, multilevel modeling can be applied to both normally and non-normally distributed outcome variables. Unless otherwise specified, all analyses presented in the following paragraphs pertain to data collected during each phase (pretest, calibration phase, and posttest).

The current study had four primary hypotheses. The first three hypotheses are contingent upon a significant three-way interaction involving trial number moderated by phase and avatar type. The three following hypotheses involve simple effects testing the form of the three-way interaction. First, based on previous findings, in terms of absolute error as the primary dependent variable (estimated distance – target distance), it was predicted that calibration to an altered avatar would occur but it would not be instantaneous as evidenced by the effect of trial number being moderated by condition in the calibration phase resulting in a steeper negative slope in the altered avatar condition than the normal avatar condition. Secondly, calibration to an altered avatar would occur more quickly than reversion back to a stored body schema as evidenced by trial moderated by phase and a steeper slope in the calibration phase than the posttest for the altered avatar condition. And thirdly, reversion back to a stored body schema would still occur in the posttest as indicated by the effect of trial number being moderated by condition in the posttest resulting in a steeper negative slope in the altered avatar condition than the normal avatar condition.

A multilevel model with absolute error as the outcome was conducted. Avatar type, phase, and trial number were entered into the model as predictors, as well as all appropriate interactions. The L1 variables of phase and trial number both had significant random effects, but their interaction did not. The presence of significant random effects indicates that there were individual differences for the effect of phase and trial number when predicting absolute error. Phase ($F(2, 13.57) = 6.14$, $p = 0.013$) and avatar type ($F(1, 22.42) = 7.84$, $p = 0.010$) had a significant main effects. The two-way interactions of phase by trial number ($F(1, 2830.66) = 18.06$, $p < 0.001$), avatar type by phase ($F(2, 26.80) = 4.53$, $p = 0.02$), and avatar type by trial number ($F(1, 25.03) = 6.14$, $p = 0.02$) were statistically significant. The three-way interaction between avatar type, phase, and trial number was statistically significant as well ($F(2, 2830.66) = 14.85$, $p < 0.001$). See Table 2. The significant three-way interaction of trial number moderated by phase and avatar type means that across the three phases, the two avatar types demonstrated different absolute error trends across trial number.

Table 2

F values, Significance Tests, and $R^2\Delta$ for Absolute Error in Experiment 1.

					R^2		
Predictors	*F*	*df*	*df*	*p*-value	Level 1	Level 2	INT
Intercept	138.39	1	20.29	<0.001	NA	NA	NA
Phase	6.14	2	13.57	0.013	2.5	NA	NA
Trial Number	0.37	1	22.33	0.550	0.03	NA	NA
Avatar Type	7.84	1	22.42	0.010	NA	30.6	NA
Phase*Trial Number	18.06	2	2830.66	<0.001	NA	NA	1.3
Avatar Type *Phase	4.53	2	26.80	0.020	NA	NA	0.02
Avatar Type *Trial Number	6.14	1	25.03	0.020	NA	NA	18.35
Avatar Type *Phase*Trial Number	14.85	2	2830.66	<0.001	NA	NA	*

TOTAL R^2	--	--	--	--	2.53	30.6	19.67

Note. *This three-way interaction is affecting error variance across multiple sources and there is not a standard practice for assessing the effect size of a L1xL1xL2 interaction.

To further investigate the significant three-way interaction of avatar type, phase, and trial number, the date file was split by phase, and three two-way interactions between avatar type and trial number were analyzed separately for each phase. The two-way interaction between avatar type and trial number was significant in the calibration phase ($F(1, 22.41) = 4.59$, $p = 0.043$) and the posttest ($F(1, 21.28) = 8.06$, $p = 0.01$), but was not significant in the pretest.

Thus, participants in the altered avatar condition exhibited greater amounts of error in their reaches across the course of trials within the calibration phase and the posttest than participants in the normal avatar group. This finding indicates that participants in the altered avatar group in the calibration phase and posttest demonstrated a greater disparity between the target distance and their reach distance across trial number, suggesting that the process of calibration to an altered avatar and the process of reversion back to the body schema occurred over the course of dozens of trials. See Table 3.

Table 3
*Predicted Means and Standard Errors for the Avatar Type*Phase*Trial Number interaction.*

Avatar Type	Pre-Test	Calibration	Post-Test
	Mean (*SE*)	Mean (*SE*)	Mean (*SE*)
Altered Avatar	5.96 (0.59)	4.48 (0.50)	4.85 (0.50)
Normal Avatar	4.39 (0.74)	2.28 (0.61)	3.51 (0.58)

Note. Presented values represent absolute error in cm when trial number is held constant at the mean.

The first hypothesis was that calibration to an altered avatar would occur but it would not be instantaneous, as evidenced by the effect of trial number being moderated by avatar type in the calibration phase resulting in a steeper negative slope in the altered avatar condition than the normal avatar condition, was supported. When investigating the significant three-way interaction of phase by avatar type by trial number, the two-way interaction of avatar type by trial number in the calibration phase was significant. As hypothesized, participants in the normal avatar group exhibited a less steep slope of trial number predicting absolute error. Participants in the altered avatar group exhibited a negative linear slope predicting absolute error across trial number in the calibration phase as predicted. Per each unit increase in trial in the calibration phase, participants in the altered avatar group exhibited a slope of -0.083 which indicates the hypothesized direction for the simple slope. This slope was significantly different than zero (t (19) = -5.018, p = <0.001, which was greater than the critical t value of -2.09). See Figure 5.

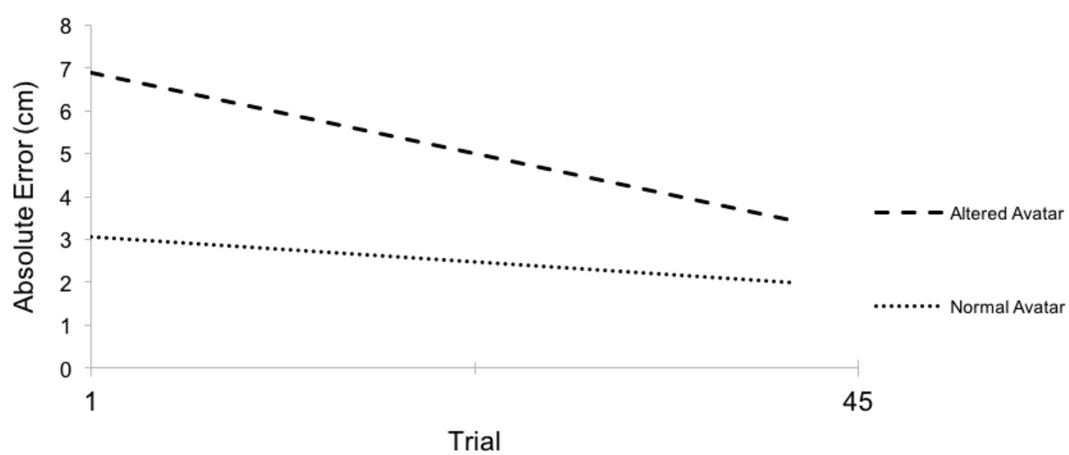

Figure 5. Simple slopes for each avatar type of trial number predicting absolute error in the calibration phase.

The second hypothesis stated that calibration to an altered avatar would occur more quickly than reversion back to a stored body schema as evidenced by trial number being moderated by phase and a steeper slope in the calibration phase than the posttest for the altered avatar condition. Support for this hypothesis was obtained because of the form of the significant interaction between avatar type, phase, and trial number. First the data file was split by avatar type, and then again by phase, to highlight the effect of trial number for each avatar type in each phase. This analysis revealed a steeper negative slope for participants in the altered avatar group in the calibration phase (coefficient = -0.083) than in the posttest phase (coefficient = -0.028). This finding indicates that calibration to an altered avatar and reversion back to a normal avatar both occurred, but calibration to an altered avatar occurred more quickly than reversion back to a normal avatar due to the steeper slope in the calibration phase.

The third hypothesis, that reversion back to a stored body schema would still occur in the posttest, was partially supported as participants in the altered avatar condition in the posttest exhibited a negative slope predicting absolute error across trials (coefficient of -0.028), but it was not significantly different than zero (t (19) = -1.803, p = 0.087, which was not greater than the critical t of -2.09). This means that in the posttest, participants in the altered avatar group exhibited decreasing amounts of absolute error over the course of the phase, suggesting that they were slowly reverting back to acting based off of their stored body schema instead of the body schema of the altered avatar they had calibrated to in the previous phase. Participants in the normal avatar condition exhibited an increase in absolute error across trials (coefficient of 0.04), which was significantly different than zero (t (19) = 2.19, p = 0.04, which was greater than the critical t = 2.09). See Figure 6. This means that in the posttest, participants in the normal avatar group exhibited increasing amounts of absolute error over the course of the phase, suggesting a decrement in performance perhaps due to fatigue from completing the same task more than 150 times. However, it should be noted that across all phases, participants in the normal avatar condition exhibited minimal amounts of error and the posttest is the only phase where those participants exhibited an increase in absolute error over trial number that was significantly different than zero. Evidence for reversion can be seen in that participants in the altered avatar group exhibited absolute error that was similar to the absolute error demonstrated by participants in the normal avatar condition at the end of the block of trials. Importantly, this effect was not immediate, as reversion only occurred after many trials had occurred in the posttest.

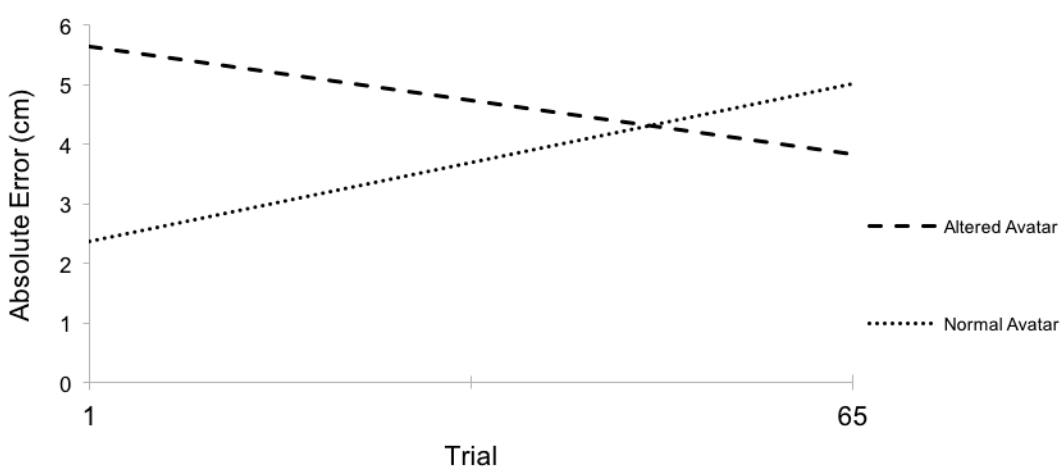

Figure 6. Simple slopes for each avatar type of trial number predicting absolute error in the posttest phase.

To investigate the fourth hypothesis, that in the posttest participants in the altered avatar group would exhibit greater under-reaches and would reach to more unreachable targets than participants in the normal avatar group, two multilevel models were run. The first investigated absolute error as the dependent variable, and included phase, avatar type, and error direction as predictors. Phase ($F(2, 2934.725) = 41.307$, $p < 0.001$), error direction ($F(1, 2939.666) = 9.369$, $p = 0.002$), and avatar type ($F(1, 23.369) = 7.257$, $p = 0.013$) all had significant main effects. The two-way interactions of phase by error direction ($F(2, 2922.107) = 8.476$, $p < 0.001$) and phase by avatar type ($F(2, 2926.347) = 5.586$, $p = 0.018$) were significant as well. The three-way interaction of phase by avatar type by error direction was also significant ($F(2, 2922.107) = 4.250$, $p = 0.014$).

Table 4
F values, Significance Tests, and $R^2\Delta$ for Absolute Error.

					R^2		
Predictors	*F*	*df*	*df*	*p*-value	Level 1	Level 2	INT
Intercept	196.744	1	24.118	<0.001	NA	NA	NA
Phase	41.307	2	2934.725	<0.001	3.2	NA	NA
Error Direction	9.369	1	2939.666	0.002	0.3	NA	NA
Avatar Type	7.257	1	23.369	0.013	NA	30.1	NA
Phase*Error Direction	8.476	2	2922.11	<0.001	NA	NA	0.5
Phase*Avatar Type	1.985	2	2918.89	0.138	NA	NA	<0.001
Avatar Type*Error Direction	5.586	1	2926.35	0.018	NA	NA	5.3
Phase*Avatar Type*Error Direction	4.250	2	2922.11	0.014	NA	NA	*
TOTAL R^2	--	--	--	--	3.5	30.1	5.3

Note. *This three-way interaction is affecting error variance across multiple sources and there is not a standard practice for assessing the effect size of a L1xL1xL2 interaction.

To further investigate the significant three-way interaction, the data file was split by phase and three two-way interactions between avatar type and error direction were conducted (one for each phase). The interactions between avatar type and error direction were significant for the pretest ($F(1, 911.509) = 10.352$, $p = 0.001$) and the posttest ($F(1, 1089.489) = 19.077$, $p < 0.001$). The interaction between avatar type and error direction in the calibration phase was not significant. Means for this interaction can be seen in Table 5 below. Overall, the difference between avatar type for under-reaching and for over-reaching differed in the pretest and posttest.

Table 5
*Predicted Means and Standard Errors for the Avatar Type*Phase*Over/Under-reach Interaction.*

	Avatar Type	Pre-Test		Calibration		Post-Test	
		Mean	*SE*	Mean	*SE*	Mean	*SE*
Under-reach	Altered	5.80	0.60	5.43	0.49	5.96	0.47
	Normal	3.16	0.75	2.98	0.61	3.12	0.67
Over-reach	Altered	6.47	0.45	4.72	0.46	4.01	0.46
	Normal	4.60	0.51	2.13	0.53	3.44	0.51

Note. Presented numbers represent absolute error in cm.

To further examine the simple effects for the significant two-way interaction between error direction and avatar type in the pretest and posttest, the effect of avatar type was examined within the pretest and posttest separately. The data file was further broken down by type of error (either an under- or over-reach). Condition was not a significant predictor for absolute error when participants under-reached in the pretest ($F(1, 15.307) = 2.457$, $p = 0.137$) or over-reached in the pretest ($F(1, 20.616) = 1.607$, $p = 0.219$). Similarly, condition was not a significant predictor of absolute error when participants over-reached in the posttest ($F(1, 21.660) = 0.018$, $p = 0.896$). Together, these findings mean that there was no significant difference between the conditions in predicting the amount of absolute error when participants over or under-reached in the pretest, and when they over-reached in the posttest. However, condition was a significant predictor for absolute error when participants under-reached in the posttest ($F(1, 28.972) = 9.596$, $p = 0.004$). According to the predicted means generated by the model, participants in the altered avatar condition ($M = 5.77$, $SE = 0.60$) under-reached significantly more than those in the normal avatar condition ($M = 2.26$, $SE = 0.96$). This finding provides further evidence that calibration to an altered avatar in the calibration phase carried over to the posttest in that when participants in the altered avatar group under-reached their reach to a target they did so by a margin significantly greater than participants in the normal avatar group. This finding suggests that calibrating to an altered avatar with an increased reaching capability in the calibration phase influenced participants to under-reach target distances to a greater extent when acting with an avatar that faithfully represented the dimensions of their own body in the posttest.

The second multilevel model was a binary logistic model that investigated correct judgment as the dependent variable, with phase, avatar type, and trial number as predictors. In terms of predicting if participants made a correct judgment, there was a significant two-way interaction of phase moderated by condition. See Table 6.

Table 6
Fixed Coefficients for the Binary Logistic Regression on Correct Judgment.

Fixed Effects		
Predictors	Coefficient (*SE*)	t
Intercept	-2.458 (0.321)	-7.668***
Phase	-0.842 (0.502)	-1.675
Trial Number	-0.01 (0.009)	-1.432
Avatar Type	0.109 (0.427)	0.427
Phase*Trial Number	0.018 (0.014)	1.312

Phase*Avatar Type	1.323 (0.599)	2.211*
Avatar Type*Trial Number	0.001 (0.011)	0.018
Phase*Avatar Type*Trial Number	-0.01 (0.016)	-0.781

*$p < 0.05$, **$p < 0.01$, *** $p < 0.001$

Overall, participants in the altered avatar condition were more likely to make incorrect judgments in the post-test (a probability of 0.114) as compared to participants in the normal avatar condition (probability of 0.041). That is, the participants in the altered avatar condition were more likely to either reach to targets that were unreachable or fail to reach to targets that were within reach. This finding also suggests that calibration to an altered avatar in the intervening calibration phase carried over to the posttest, in that participants in the altered avatar condition continued to reach to target distances that would have been reachable in the calibration phase with the altered avatar but were no longer reachable in the posttest. See Table 7.

Table 7
Predicted Probability of Making an Incorrect Reach Judgment.

Condition	Pre-Test		Post-Test	
	Mean	SD	Mean	SD
Altered Avatar	0.065	0.028	0.114	0.037
Normal Avatar	0.057	0.016	0.041	0.009

In summation, the analysis of Experiment 1 yielded interesting and novel findings. As expected, when predicting absolute error from trial number, phase, and avatar type, there was a significant three-way interaction of trial number moderated by phase and avatar type. The specific form of this three-way interaction was shown by the fact that calibration to an altered avatar occurred in the presence of explicit feedback. Calibration was evidenced by the direction of the simple slope for trial number predicting absolute error in the calibration phase for participants in the altered avatar group and by the carryover effects demonstrated in the posttest by the altered avatar group (i.e., quite large absolute error at the beginning of the posttest). Then, in the absence of explicit informative feedback in the posttest, reversion back to a stored body schema occurred as well. Evidence for reversion can be seen in that participants in the altered avatar group exhibited absolute error that was similar to the absolute error demonstrated by participants in the normal avatar condition at the end of the block of trials. This effect was not immediate, as reversion only occurred after many trials had occurred in the posttest. Interestingly, calibration to an altered avatar occurred more quickly than reversion back to a normal avatar due to the steeper slope in the calibration phase. Lastly, participants in the altered avatar group exhibited greater under-reaches than participants in the normal avatar group in the posttest *and* made more incorrect judgments in the posttest.

Taken together, all of these results confirm that actors can calibrate to an avatar with different bodily dimensions than their own, and that this process of calibration is not instantaneous. Then, once an actor has calibrated to an avatar with different bodily dimensions than their own physical body, in the absence of explicit feedback regarding their reaching behavior in the virtual world participants begin to revert back to their stored body schema.

CHAPTER FOUR

EXPERIMENT TWO

VR simulations are often used for training purposes. VR is ideally suited as a training medium because training programs can be implemented where they cannot be in the real world when it is excessively dangerous, too expensive, or too difficult to control the training scenario (Rose, Attree, Brooks, Parslow, & Penn, 2000). Virtual training programs allow the administrator complete control over presentation of stimuli and the type of feedback that the trainee receives. VR training applications have been implemented in a variety of settings ranging from training airline pilots (Lintern, Roscoe, Koonce, & Segal, 1990), firefighters (Bliss, Tidwell, & Guest, 1997), police officers (Bertram, Moskaliuk, & Cress, 2015), and surgeons (Hyltander, Liljegren, Rhodin, & Lonroth, 2002).

To assess the carry over effects from one modality (VR) to another (the real world), a transfer of calibration paradigm will be used. Transfer of training paradigms utilize a pretest, exposure, posttest design. Transfer is considered to have occurred if the calibration that occurred in an exposure phase carries over to the posttest. The issue of transfer of calibration between training in virtual environments and performance in the real world is crucial to consider, as studies have come to different conclusions.

It has been assumed that training in VR will transfer to real world performance, and there is conflicting evidence to support this claim. For example, the early findings from Kozak, Hancock, Arthur, and Chrysler (1993) suggested that transfer from virtual training to the real world might not occur. When we act in the real world, we act as a unitary system that integrates information from a variety of perceptual systems (kinesthetic, haptic, visual, etc.). In the real world, the haptic perceptual system provides the actor with information regarding limb position and movement, in addition to information regarding wielded object properties like positon, orientation, weight, etc. (Pagano and Turvey, 1992, 1998). Yet in the virtual world, there is often an interruption between the information specifying kinesthetic and visual invariants. As such, training on tasks in VR may not transfer to the real world due to this interruption in specifying pertinent information.

Previous research has also shown differences in performance on tasks completed in the real world compared to the same task completed in VR (Napieralski et al., 2011). Ebrahimi et al. (2016) showed that participants were more accurate performing a reaching task in the real world compared to the virtual world. Bufton, Campbell, Howie, and Straker (2014) showed that when playing the same game (ping-pong) in the real world or in a virtual setting, participants exhibited different movement patterns across the two modalities. Their findings suggest that the difference in movement patterns may interfere with learning the real-world motor skill.

Other studies have found evidence to suggest that training in VR does transfer to real world environments (Bertram et al., 2015; Ganier, Hoareau, & Tisseau, 2014; Hyltander et al., 2002; Larrue et al., 2014; Regian, 1997; Rose et al., 2000). Interestingly, some training programs do not represent the user with an avatar, while others use very low fidelity avatars (Koritnik, Koenig, Bajd, Riener, & Munih, 2010), avatars that are not scaled to the dimensions of the user (Bertram et al., 2015; Bufton et al., 2014), or disembodied avatar limbs (Ganier et al., 2014; Grabowski and Jankowski, 2015). Thus, it is unknown if the size of the avatar used to represent the user has an impact on the transfer of calibration of action capabilities to the real world. Must the avatar be scaled exactly to the dimensions of the user's biological body for the skills learned in VR to transfer to the real world?

As recommended by Ebrahimi et al. (2014) and Bufton et al. (2014), this study will test whether calibration in VR carries over to performance in the real world, and if the size of the avatar impacts the transfer of calibration. Experiment 2 will replicate the first proposed experiment with one crucial change: instead of having all participants in all conditions complete each block of trials in virtual reality, the second experiment will have participants perform the same reaching task in the real world in the pre- and post-tests. In this way, we can compare the data obtained in a study that incorporates acting in the real world and virtual world to a study done only in VR.

Hypotheses

It will be important to compare the results of the two experiments. A comparison between the pretest from Experiment 1 and the pretest in Experiment 2 will be conducted to investigate how performance differs between reaching in the real world and reaching in a virtual environment with an avatar arm. We will also compare the posttest in Experiment 1 to the posttest in Experiment 2 to see differences in how calibration persists in VR compared to training in VR and then switching bac to acting in the real world.

The current study has three primary hypotheses. Based on previous findings, we predict that calibration to a faithful avatar in the experimental calibration phase will occur more quickly than calibration to an altered avatar. This means that participants in the normal avatar condition will have a smaller intercept and a much flatter slope than participants in the altered avatar condition in the experimental calibration phase who will have a large intercept and steep negative slope. Further, we predict that reversion back to one's normal body capabilities in the posttest will occur more quickly in the normal avatar condition as compared to the altered avatar condition. This means that participants in the normal condition will have a smaller intercept and flatter slope than participants in the altered condition in the posttest, who will have a large negative intercept and a positive slope. Again, we predict that reversion back to the user's normal body representation will still occur in the posttest, and reversion will be evidenced by participants in the altered avatar condition demonstrating the same amount of absolute error in the posttest as participants in the normal avatar condition.

CHAPTER FIVE

EXPERIMENT TWO METHOD

Participants

Twenty-three undergraduate students (15 females and 8 males, *M* = 19.28, *SD* = 1.1) from Clemson University participated in this experiment. Reaching data from one participant was discarded for failure to follow directions. Participants were required to be right handed as all equipment used was for right-handed participants. All participants received credit in their psychology courses in exchange for participation. As participants entered the testing area, they were given a brief overview of the purpose of the experiment and informed consent was obtained. Participants were randomly assigned to either the altered avatar condition or normal avatar condition.

Design

The second experiment utilized a 2 (Avatar Type: Altered avatar vs. Normal Avatar) by 3 (Phase: PreTest, Calibration, PostTest) mixed groups design. Avatar type was a between subject variable and phase was a within subject variable. The normal avatar condition involved use of an avatar's arm that was directly proportional to the dimensions of the user's own arm. The altered avatar condition involved use of an avatar whose arm length, and thus reaching capabilities, were increased by 30 cm.

Materials and Apparatus

The materials and apparatus used were the same as in Experiment 1. Regardless of modality, all participants completed reaches while equipped with the Vive controller.

For the trials completed in the real-world participants reached for a target area that was the same size as the virtually presented target.

Procedure

The procedure was the same as Experiment 1, except all participants completed their reaches in the real world in the pretest and posttest while the calibration phase was completed in VR. Participants in the normal avatar condition and altered avatar condition reached with a faithful or altered avatar in VR in the calibration phase, respectively.

CHAPTER SIX

EXPERIMENT TWO RESULTS

Body Ownership

Mean responses on each of the six items in the body ownership questionnaire were compared between groups. As can be seen in Table 8 there were no significant differences in feelings of body ownership between participants in the normal tool and long tool groups.

Table 8

Means, Standard Deviations, Standard Errors, and Significance Values for responses to the Body Ownership Questionnaire in Experiment 2.

	Condition	N	Mean	Std. Deviation	Std. Error Mean	Significance
Question1	Normal	12	6.8333	1.89896	.54818	.984
	Long	11	6.8182	1.53741	.46355	
Question2	Normal	12	6.5000	1.83402	.52944	.957
	Long	11	6.4545	2.16165	.65176	
Question3	Normal	12	8.7500	.96531	.27866	.967
	Long	11	8.7273	1.61808	.48787	
Question4	Normal	12	3.8333	2.40580	.69449	.915
	Long	11	3.7273	2.28433	.68875	
Question5	Normal	12	3.2500	2.22077	.64108	.714
	Long	11	3.6364	2.76668	.83419	
Question6	Normal	12	6.7500	2.76751	.79891	.947
	Long	11	6.8182	1.94001	.58493	

In response to the post-data collection manipulation check, 17 participants (74% of total) indicated that they noticed anything odd during the course of the experiment while six participants (26 % of total) indicated that nothing seemed odd. Of those 17 participants who responded that they noticed something odd during the experiment, only five participants (29% of yes responders) mentioned anything about the arm of the avatar being extended, manipulated, or larger. Common responses were, ‘was a little shaky at times’, ‘skin looked weird’, or ‘arms were long’. Broken down by avatar type, of the 12 participants in the normal avatar group, nine participants (75%) indicated they noticed something odd and three participants (25%) said they did not notice anything. Of the 11 participants in the altered avatar group, eight participants (73%) indicated they noticed something odd and three participants (27%) said they did not notice anything. Of those 11, five participants (45%) specifically mentioned something about the arm of the avatar being extended, manipulated, or larger.

Outlier Analysis

For each analysis, individual outlier analyses for full models were conducted. Residuals were obtained, standardized, and examined for any potential outliers that were outside of the normal distribution (Cohen et al., 2003). Outlier analysis was based on data visualization as well. Data points that were likely due to malfunctions in the tracking equipment and were not physically possible were removed for each specific analysis. In all of the analyses less than 2% of the trials were removed due to outliers.

Hierarchical Linear Modeling

Figure 7 demonstrates the raw data in terms of estimated distance (the distance to which participants reached with the tip of the tool) and presented target distance. The overall data is shown, as well as the data for each phase.

The intraclass correlation (ICC) of the intercept only model (null model) was used to assess the overall nesting within participants for each of the main dependent variables (correct judgment and absolute error). Due to the repeated-measures design of the experiment, variables had significant nesting within participants. For example, the obtained ICC for absolute error as the DV was approximately 7%.

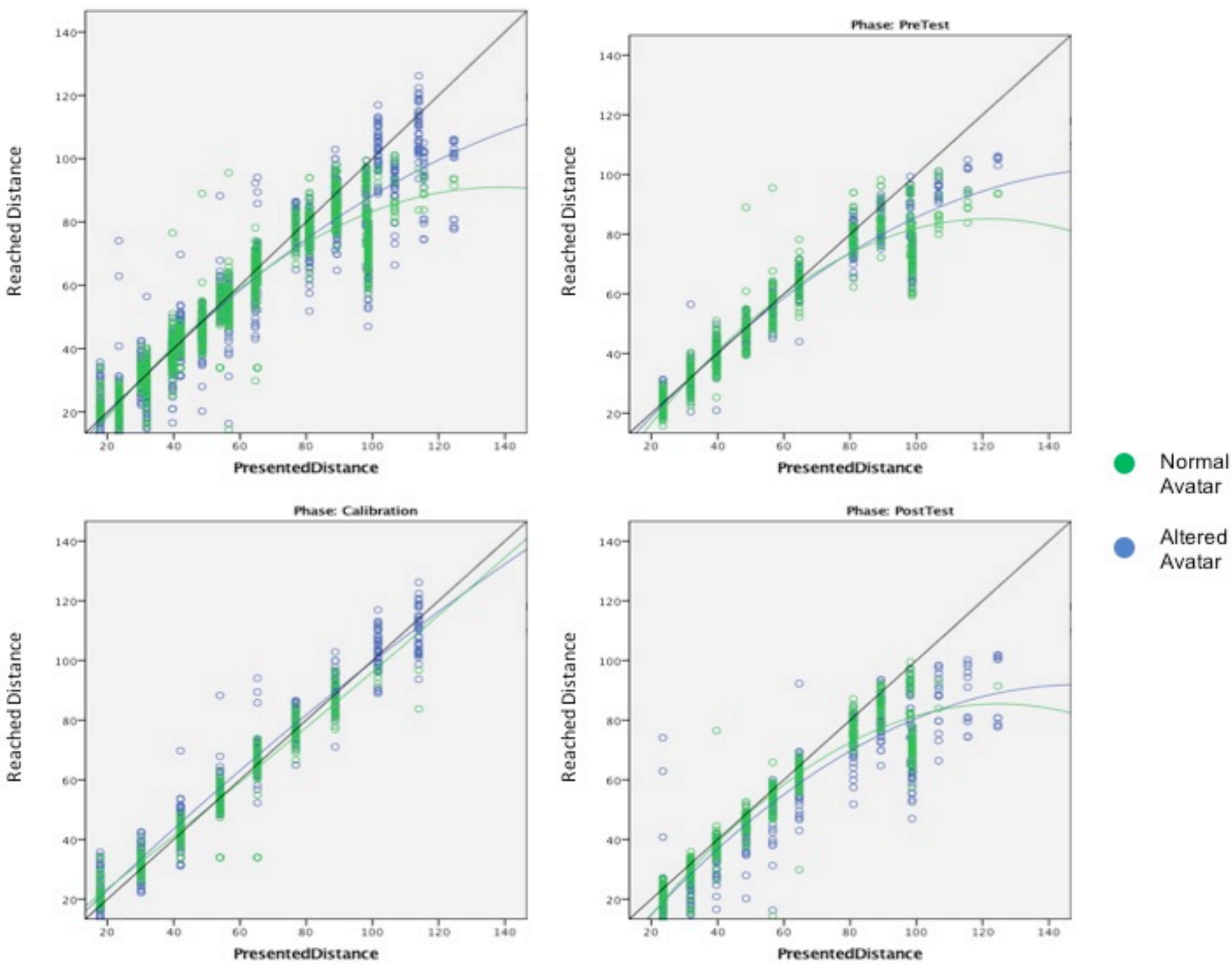

Figure 7. Estimated distance as a function of presented distance (clockwise from top left) a) overall b) pre-test c) calibration phase d) post-test. The solid black line in each graph represents perfect performance (y=1x+0).

The current study has three primary hypotheses, all of which are contingent upon an interaction of trial number moderated by avatar type and trial number. Based on previous findings, we predicted that calibration to a faithful avatar in the calibration phase would occur more quickly than calibration to an altered avatar. This means that when predicting absolute error from trial number participants in the normal avatar condition will have a smaller intercept and a much flatter slope than participants in the altered avatar condition in the experimental calibration phase who will have a large intercept and steep negative slope. Further, we predicted that reversion back to one's normal body capabilities in the posttest would occur more quickly in the normal avatar condition as compared to the altered avatar condition. This means that participants in the normal condition would have a smaller intercept and flatter slope than participants in the altered condition in the posttest, who will have a large negative intercept and a positive slope. Again, we predict that reversion back to the user's normal body representation would still occur in the posttest, and reversion would be evidenced by participants in the altered avatar condition demonstrating the same amount of absolute error in the posttest as participants in the normal avatar condition.

A multilevel model with absolute error as the outcome was conducted. Avatar type, phase, and trial number were entered into the model as predictors, as well as all appropriate interactions. Phase ($F(2, 2696.697) = 47.524$, $p < 0.001$) and trial number ($F(1, 2692.898) = 5.486$, $p = 0.019$) had significant main effects. The two-way interactions of phase by trial number ($F(2, 2685.720) = 3.803$, $p = 0.022$), avatar type by phase ($F(2, 2687.201) = 4.148$, $p = 0.016$), and avatar type by trial number (F(2, 2685.761) = 6.458, *p* = 0.011) were all statistically significant. The three-way interaction between avatar, phase, and trial number was not statistically significant (see Table 9). This means that in each phase, participants in the two conditions demonstrated similar absolute error across all trials, suggesting that they calibrated to either the tool or the avatar they were using at similar rates. However, the significant two-way interactions between phase and trial number indicates there was a difference in the rate of calibration over each of the three phases. The significant two-way interaction between phase and avatar type suggests that there was a difference in the mean absolute error demonstrated between the two conditions across each phase. The significant two-way interaction between condition and trial number indicates that the altered avatar group and normal avatar group differed in the mean absolute error demonstrated across trials in general.

Table 9
F values, Significance Tests, and $R^2\Delta$ for Absolute Error in Experiment 2.

					R^2		
Predictors	*F*	*df*	*df*	*p*-value	Level 1	Level 2	INT
Intercept	161.080	1	29.747	<0.001	NA	NA	NA
Phase	47.524	2	2696.697	<0.001	3.3	NA	NA
Trial Number	5.486	1	2692.898	0.019	0.2	NA	NA
Avatar Type	1.219	1	19.411	0.283	NA	1.3	NA
Phase*Trial Number	3.803	2	2685.720	0.022	NA	NA	0.2
Avatar Type *Phase	4.148	2	2687.201	0.016	NA	NA	5.2
Avatar Type *Trial Number	6.458	1	2685.761	0.011	NA	NA	1.4
Avatar Type *Phase*Trial Number	0.032	2	2685.720	0.968	NA	NA	*
TOTAL R^2	--	--	--	--	3.5	1.3	6.8

Note. *This three-way interaction is affecting error variance across multiple sources and there is not a standard practice for assessing the effect size of a L1xL1xL2 interaction.

Overall, across all phases participants in the normal avatar condition exhibited less absolute error (M = 7.68, SE = 0.664) than participants in the altered avatar condition (M = 10.00, SE = 0.686). To further investigate the significant two-way interaction between phase and avatar type, means were produced for each condition in each phase. The means for each avatar type in each phase can be seen in the table below. In general, there was no difference in absolute error between avatar types in the pretest and calibration phases. However, participants in the altered avatar condition exhibited greater absolute error in the posttest (M = 10.26, SE = 0.80) than participants in the normal avatar condition (M = 6.77, SE = 0.77).

Table 10
Mean Absolute Error for each Condition broken down by Phase.

Avatar Type	Pre-Test	Calibration	Post-Test
	Mean (*SE*)	Mean (*SE*)	Mean (*SE*)
Altered Avatar	6.79 (0.80)	4.62 (0.84)	10.26 (0.80)
Normal Avatar	7.56 (0.77)	4.18 (0.85)	6.77 (0.77)

Note. Presented values represent absolute error in cm.

A graph illustrating the significant two-way interactions of trial number moderated by condition can be seen in Figure 8. As can be seen in the graph, participants in the altered avatar condition showed decreasing amounts of absolute error over trials. After splitting the file by avatar type, to investigate simple effects, it was revealed that the simple slopes for the altered avatar and normal avatar groups were not significantly different from zero.

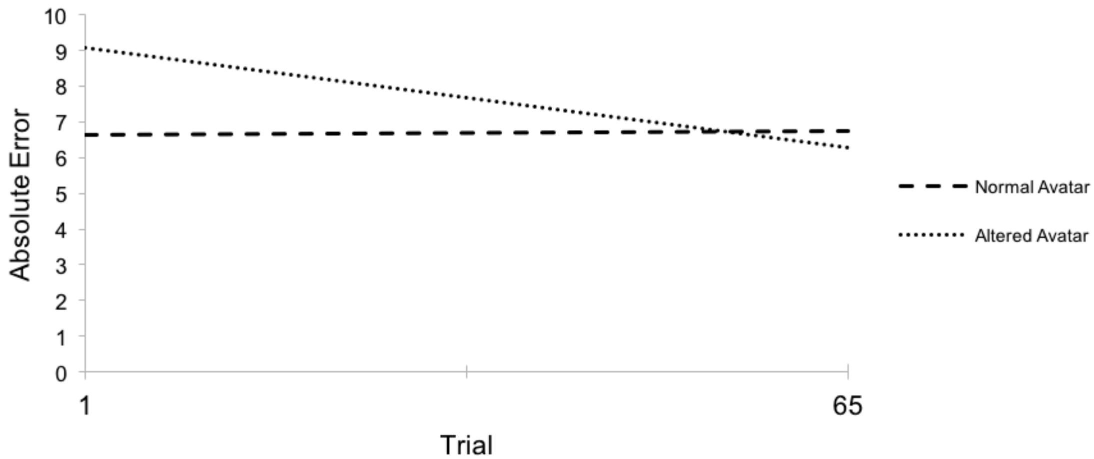

Figure 8. Two-way interactions of trial number moderated by condition.

Lastly, simple effects were identified for the phase by trial number interaction. The data file was split by phase to examine the effect of trial number in each phase. Trial number was not a significant predictor of absolute error in the pretest. However, trial number was a significant predictor of absolute error in the calibration phase ($F(2, 731.421) = 23.056$, $p < 0.001$) and the posttest ($F(2, 975.929) = 6.428$, $p = 0.011$). Across both avatar types in the calibration phase, per each unit increase in trial number, participants exhibited a decrease in absolute error of -0.06 cm. Similarly, across both avatar types in the calibration phase, per each unit increase in trial number, participants exhibited a decrease in absolute error of -.04 cm.

Due to the non-significant three-way interaction of trial number moderated by phase and condition, none of the three hypotheses were fully supported.

To replicate the analysis performed on Experiment 1, additional multilevel models were run on Experiment 2 data. The first investigated absolute error as the dependent variable, and included phase, avatar type, and error direction as predictors. Phase ($F(2, 2703.573) = 11.354$, $p < 0.001$) and error direction ($F(1, 2660.361) = 86.752$, $p < 0.001$) had significant main effects. The two-way interactions of phase by error direction ($F(2, 2703.017) = 22.946$, $p < 0.001$), phase by avatar type ($F(2, 2698.417) = 9.364$, $p < 0.001$), and avatar type by error direction ($F(2, 2661.239) = 9.008$, $p = 0.003$) were all significant as well. The three-way interaction of phase by avatar type by error direction was not significant ($F(2, 2703.017) = 2.294$, $p = 0.101$).

Table 11
F values, Significance Tests, and $R^2\Delta$ for Absolute Error regarding Under and Over-reaches.

					R^2		
Predictors	*F*	*df*	*df*	*p*-value	Level 1	Level 2	INT
Intercept	183.673	1	19.835	<0.001	NA	NA	NA
Phase	11.354	2	2703.573	<0.001	0.7	NA	NA
Error Direction	86.752	1	2660.361	<0.001	2.8	NA	NA
Avatar Type	1.280	1	19.403	0.272	NA	1.6	NA
Phase*Error Direction	22.946	2	2703.017	<0.001	NA	NA	<0.001
Phase*Avatar Type	9.364	2	2698.417	<0.001	NA	NA	3.8
Avatar Type*Error Direction	9.008	1	2661.239	0.003	NA	NA	56.2
Phase*Avatar Type*Error Direction	2.294	2	2703.017	0.101	NA	NA	*
TOTAL R^2	--	--	--	--	3.5	1.6	60.0

Note. *This three-way interaction is affecting error variance across multiple sources and there is not a standard practice for assessing the effect size of a L1xL1xL2 interaction.

Means for each of the two-way interactions can be seen in Table 12 below. As indicated in the interaction of error direction moderated by phase, participants under-reached their reaches to targets to a greater extent in the pretest and posttest as compared to the calibration phase. Next, in the posttest only, participants in the altered avatar condition exhibited larger amounts of absolute error than participants in the normal avatar condition. Lastly, regardless of phase, participants in both avatar types exhibited similar amounts of error when under-reaching, but participants in the altered avatar condition exhibited greater over-reaches than participants in the normal avatar condition.

Table 12
Predicted Means and Standard Errors for the Phase by Error Direction (top), Phase by Avatar Type (middle) and Avatar Type by Error Direction (bottom) Interactions.

Error Direction	Pre-Test		Calibration		Post-Test	
	Mean	*SE*	Mean	*SE*	Mean	*SE*
Under-reach	8.82	0.54	4.97	0.65	9.40	0.52
Over-reach	3.90	0.64	4.70	0.57	3.39	0.80

Avatar Type	Pre-Test		Calibration		Post-Test	
	Mean	*SE*	Mean	*SE*	Mean	*SE*
Altered	6.05	0.74	4.91	0.75	7.97	0.84
Normal	6.67	0.71	4.76	0.76	4.81	0.75

Error Direction	Avatar Type		
		Mean	*SE*
Under-reach	Altered	7.60	0.70
	Normal	7.85	0.68
Over-reach	Altered	5.01	0.78
	Normal	2.98	0.73

Note. Presented numbers represent absolute error in cm.

To further examine the simple effects for the significant two-way interactions, the effect of avatar type was examined for each phase and for over/under-reaches separately. The data file was broken down by phase and by type of error (either an under- or over-reach), respectively. Condition was not a significant predictor of absolute error for any of the phases. Further, condition was not a significant predictor of absolute error when participants under-reached. However, condition was a significant predictor of absolute error when participants over-reached ($F(1, 16.717) = 5.313$, $p = 0.034$). According to the predicted means generated by the model (and not accounting for the effect of trial number), participants in the altered avatar condition ($M = 4.82$, $SE = 0.50$) over-reached significantly more than those in the normal avatar condition ($M = 3.21$, $SE = 0.49$). Lastly, the data file was again split by phase to investigate the effect of error direction. Error direction was a significant predictor in the pretest ($F(1, 664.850) = 88.905$, $p < 0.001$) and the posttest ($F(1, 994.485) = 34.512$, $p < 0.001$), but not in the calibration phase. According to the predicted means generated by this model, regardless of avatar type, participants under-reached their reaches to targets to a greater degree in the pretest ($M = 9.14$, $SE = 0.53$) than their over-reaches ($M = 3.22$, $SE = 0.64$). The same pattern was found in the posttest, where participants under-reached their reaches to targets to a greater degree ($M = 9.01$, $SE = 0.94$) than their over-reaches ($M = 3.96$, $SE = 1.18$). This finding suggests that when participants were acting and receiving no feedback they tended to under-reach their reaches to targets, and that in the presence of no feedback participants reverted back to acting with their stored body schema. Further, these findings suggest that participants had no difficulty calibrating to reaching with an avatar in the intervening calibration phase. This is likely to have occurred due to the presence of explicit and informative visual feedback during the calibration phase.

The second multilevel model was a binary logistic model that investigated whether participants correct judgment as the dependent variable, with phase, avatar type, and trial number as predictors. In terms of predicting if participants made a correct judgment, there was a significant two-way interaction of phase moderated by condition. See Table 13.

Table 13
Fixed Coefficients for the Binary Logistic Regression on Correct Judgment.

Fixed Effects		
Predictors	Coefficient (*SE*)	t
Intercept	-1.463 (0.236)	-6.196***
Phase	-0.022 (0.10)	-0.215
Trial Number	-0.002 (0.003)	-0.785
Avatar Type	0.151 (0.227)	0.664
Phase*Trial Number	0.001 (0.005)	0.230
Phase*Avatar Type	0.543 (0.202)	2.694**
Avatar Type*Trial Number	-0.004 (0.005)	-0.718
Phase*Avatar Type*Trial Number	0.001 (0.011)	0.097

**p* < 0.05, **p < 0.01, *** p <0.001

Overall, participants in the altered avatar condition were more likely to make incorrect judgments in the post-test (a probability of 0.215) as compared to participants in the normal avatar condition (probability of 0.149). That is, the participants in the altered avatar condition were more likely to either reach to targets that were unreachable or fail to reach to targets that were within reach. This finding also suggests that calibration to an altered avatar in the intervening calibration phase carried over to the posttest, in that participants in the altered avatar condition continued to reach to target distances that would have been reachable in the calibration phase with the altered avatar but were no longer reachable in the posttest. See Table 14.

Table 14
Predicted Probability of making an Incorrect Reach Judgment.

Condition	Pre-Test		Post-Test	
	Mean	SD	Mean	SD
Altered Avatar	0.177	0.061	0.215	0.071
Normal Avatar	0.190	0.053	0.149	0.044

Contrary to the findings of Experiment 1, when predicting absolute error there was not a significant interaction of trial number moderated by phase and avatar type. This means that the two groups did not differ in the amount of absolute error they demonstrated across the course of trials across phases. However, there was a significant interaction of phase and avatar type in predicting absolute error, where participants in the altered avatar condition exhibited more absolute error in the posttest. Further, when predicting absolute error based off phase, avatar type, and error direction, there were significant two-way interactions of phase by error direction, phase by avatar type, and avatar type by error direction. Lastly, based off the significant interaction between phase and avatar type in the binary logistic regression, it was revealed that participants in the altered avatar group made more incorrect judgments in the posttest than participants in the normal avatar group.

The last finding is very important, as it suggests that calibration to an altered avatar in VR can carry over to the real world when participants begin to act with their normal body again. Taken together, all of these findings suggest that calibration to an altered avatar and to normal avatar occur at the same rate. Then, once an actor has calibrated to an avatar in VR, but then switches back and begins to act in the real world without explicit feedback, reversion back to an actor's normal body occurs at the same rate for each condition as well. However, despite there being no difference in the rate at which calibration and reversion occur between the two avatar types, there was a difference in the mean absolute error demonstrated between the two conditions across each phase. Regardless of trial number, participants in the altered avatar condition exhibited greater discrepancies in their reaches to targets in both in the calibration phase and the posttest than participants in the normal avatar condition for each respective phase. Overall, the results suggest that calibration to an avatar can occur, and that this calibration in VR can carry over to the real world for a period of time before the process of reversion back to the body schema occurs.

Comparison of Data Between Experiment 1 and Experiment 2

Before continuing, there is the possibility that even reaching with an avatar arm that faithfully represents the dimensions of the user will result in performance that differs from reaching performance in the real world. Because of this possibility, we compared the results of the pretest in Experiment 1 (which was completed in VR) to the pretest of Experiment 2 (which was completed in the real world). Participants in the pretest of

Experiment 2, which served as our baseline comparison condition, engaged in the same pretest procedure, except all reaches were completed in the real world.

Pretest data from the first experiment was compared to pretest data from the second experiment. A multilevel model with absolute error as the dependent variable was run with experiment and trial number as predictors. The interaction between the two terms was included as well. Experiment had a significant main effect ($F(1, 97.63) = 7.368$, $p = 0.008$), but trial number was not significant (see Table 15). In the pretest only, regardless of trial number, participants who completed the pretest in the real world exhibited an average absolute error of 6.51 cm ($SE = 0.453$), while participants who completed the pretest in VR had an average absolute error of 4.19 cm ($SE = 0.423$). See Table 16. The two-way interaction of experiment by trial number was not significant. Overall, participants in the real world exhibited larger disparities when reaching to targets than participants in virtual reality, and the change in reaching errors over the course of the pretest did not differ between the two experiments.

Table 15
F values, Significance Tests, and $R^2\Delta$ for Absolute Error Pretest comparison across Experiments.

					R^2		
Predictors	*F*	*df*	*df*	*p*-value	Level 1	Level 2	INT
Intercept	202.374	1	97.63	<0.001	NA	NA	NA
Trial Number	0.022	1	1981.41	0.882	<0.001	NA	NA
Experiment	7.368	1	97.63	0.008	NA	34.6	NA
Experiment*Trial Number	0.373	1	1981.41	0.541	NA	NA	<0.001
TOTAL R^2	--	--	--	--	<0.001	34.6	<0.001

Table 16

Predicted mean Absolute Error for the Pretest across both Experiments.

Experiment Modality	Mean (*SE*)	Pairwise comparison
Real World	6.51 (0.453)	$p = 0.001$
Virtual Reality	4.19 (0.423)	

If there were no differences between Experiment 1 and Experiment 2 data collected in the pretest we could conclude that the performance of participants acting with an avatar in VR is consistent with real world performance. However, in the current comparison, acting with an avatar in VR gives an added benefit of decreased disparity between target distance and estimated distance when reaching to targets in the virtual world compared to the real world.

Additionally, posttest data from the first experiment was compared to posttest data from the second experiment. A multilevel model with absolute error as the dependent variable was run with experiment and trial number as predictors. The interaction between the two terms was included as well. Experiment had a significant main effect ($F(1, 63.986) = 13.011$, $p = 0.005$), and trial number was significant as well ($F(1, 2037.880) = 7.876$, $p = 0.001$). In the posttest only, regardless of trial number, participants who completed the posttest in the real world exhibited an average absolute error of 7.77 cm ($SE = 0.762$), while participants who completed the posttest in VR had an average absolute error of 4.21 cm ($SE = 0.713$). The two-way interaction of experiment by trial number was not significant. See Table 17. Overall, participants in the real world exhibited larger disparities when reaching to targets than participants in virtual reality. Further, there was a change in reaching errors over the course of the posttest. As trial number increased by one, absolute error decreased by .03 cm. But this change in absolute error did not differ between the two experiments, and the change in reaching errors over the course of the posttest did not differ between the two experiments. Overall, in the current comparison there seems to be a negative effect on reaching to target distances after switching modalities, meaning participants who completed the calibration phase in VR and the posttest in the real world (i.e. participants in Experiment 2) exhibited greater disparities in reaching to targets. Those participants who acted in a congruent modality for each phase showed significantly less disparity in their reaches to targets (see Table 18).

Table 17
F values, Significance Tests, and $R^2\Delta$ for Absolute Error Posttest comparison across Experiments.

					R^2		
Predictors	*F*	*df*	*df*	*p*-value	Level 1	Level 2	INT
Intercept	135.075	1	63.986	<0.001	NA	NA	NA
Trial Number	7.876	1	2037.880	0.001	0.3	NA	NA
Experiment	13.011	1	63.986	0.005	NA	25.5	NA
Experiment*Trial Number	1.467	1	2037.880	0.226	NA	NA	5.8
TOTAL R^2	--	--	--	--	0.3	25.5	5.8

Table 18
Predicted mean Absolute Error for the Posttest across both Experiments.

Experiment Modality	Mean (*SE*)	Pairwise comparison
Real World	7.77 (0.762)	*p* = 0.001
Virtual Reality	4.21 (0.713)	

CHAPTER SEVEN

DISCUSSION

Human actors are quite adept at assimilating tools designed to extend reach into their body schema, and this alters their perception of distance in reachable space (Proffitt & Linkenauger, 2013; Witt, Proffitt, & Epstein, 2005). The current studies involved a direct manipulation of arm length capitalizing on virtual reality technology. Specifically, the present studies investigated whether an actor can calibrate to the action capabilities of an avatar that possess different anthropometric dimensions than themselves. In Experiment 1, it was hypothesized that participants would be able to calibrate to an altered avatar in the presence of feedback, and that the process of calibration would occur more quickly than reversion back to the dimensions of a faithful avatar in the posttest. The results suggest that calibration to an extended avatar did occur, but that participants need a prolonged period of exposure, in this case upwards of 45 trials, to reaching with the extended avatar before they properly calibrate to the altered action capabilities. Subsequently, after calibration has occurred and explicit feedback regarding their reaching behavior is removed, the participants revert back to their stored body schema.

While Experiment 1 occurred in VR only, Experiment 2 compared performance in the real world before and after a calibration phase that occurred in VR. Again, it was hypothesized that participants would be able to calibrate to an altered avatar in the presence of feedback and that the process of calibration would occur more quickly than reversion in the posttest. This hypothesis was not supported as the rates of calibration for each avatar type were not significantly different in the various phases. Relatedly, the significant two-way interaction of phase and trial number revealed that the process of calibration occurred at a quicker rate than the process of reversion, regardless of avatar type. However, the significant two-way interaction between phase and avatar type revealed that participants in the altered avatar condition exhibited greater amounts of error than participants in the normal avatar condition in the posttest.

The present results support the idea that participants are able to calibrate to using an avatar with different anthropometric dimensions than their own body, but this process of calibration is not instantaneous. Rather, dozens of repeated trials are needed for calibration to occur. Further, the results indicate that the size of an avatar, which in the current situation is directly correlated with action capabilities, influences distance estimations in VR. This suggests that the ability to act *and* accurately perceiving one's action capabilities are vital components of virtual environments, especially if the virtual environment is to be used for learning and training that needs to transfer to the real world.

Another way to further investigate the current data set is to perform MLM analysis with a five-way interaction of phase, avatar type, trial number, error direction, and presented distance. In this way, more specific conclusions can be made regarding the behaviors exhibited by participants acting with a normal or altered avatar in the various phases. For example, information would be obtained that would indicate the direction in which participants in the altered avatar group are exhibiting reaching errors in the calibration phase and posttest.

Contributions to the Calibration Literature

The results from both experiments contribute to the existing literature regarding calibration. Calibration refers to the necessary process of scaling between aspects of the environment and action capabilities of an actor. Calibration, as distinct from the concept of affordances, relies on feedback to more accurately perceive affordances. Generally, this process allows for actors to distinguish between possible and impossible opportunities for action in a particular environment. Phrased differently, calibration is engaging in environmentally directed action that is informed by a scaling between action and perception (Mon-Williams & Bingham, 2007). Functionally speaking, calibration is vital for this scaling *and* to ensure the stability of that scaling. Due to the fact calibration is a form of measurement, and all measurements involve some level of noise, stability becomes very important for calibration to be functionally effective (Bingham & Pagano, 1998). Calibration cannot eliminate all error inherent in any behavior, but the task of the process of calibration is to effectively keep error to a minimum. The process of calibration is important because it keeps error at a minimum all while keeping the behavior relatively constant across calibration.

The mechanism of calibration entails a mapping between sensory information and information generated by action, both of which can be manipulated to cause calibration to occur (van Andel et al., 2017). Within the current study action capabilities were manipulated through the extension of an avatar arm, so participants were tasked with scaling their new action capabilities to match with embodied units of perception. Calibration was judged to have occurred by taking into account action judgments and measures regarding control of motor behavior.

Recent work has identified opportunities to improve the existing literature on calibration and its underlying mechanisms. Van Andel and colleagues (2017) state that studies explicitly studying the rate of calibration in various situations is lacking, as it is very important to identify the amount of experience that is required for effective calibration. Similarly, they call for an investigation of individual differences in the process of calibration. The current studies contribute directly to both of these existing holes in the literature.

First, the present studies revealed that while calibration to an altered avatar can occur, the process is relatively slow compared to previously published results regarding the rate of calibration as discussed below. A unique contribution of the current research is that we were able to track the change in our primary DV (absolute error) over the course of each individual trial in the calibration phase and the posttest without aggregating data. For example, out of the 23 papers included in van Andel et al.'s (2017) systematic review of calibration research, only nine included references to the rate of calibration. In this way, we are able to make specific contributions to the existing literature regarding the rate of calibration. In Experiment 1, participants in the altered avatar group needed upwards of 65 trials before their exhibited absolute error was similar to the absolute error demonstrated by the normal avatar condition. All previous work has made the claim that some amount of experience is necessary for calibration to occur, but the amount of experience that is necessary is unclear, as it most likely depends on many factors (i.e. the setting of the task, the demands of the task, experience with the task, etc.). Previous research has claimed that for certain behaviors such as braking, calibration can occur in as little as one second (Fajen, 2007), or as long as 30 minutes when judging sitting and stepping height (Mark, 1987; as described in Mark et al., 1990). Other studies claim that minimal experience, such as only a single reach, is sufficient for calibration to occur (Linkenauger, Bülthoff, & Mohler, 2015), while other studies suggest only five trials are necessary for calibration to occur in an object interception task (Scott & Gray, 2010; see also Bourgeois & Coello, 2012). In many studies, phases such as "people only needed a handful of trials of perceptual-motor feedback to recalibrate and perform perceptual motor tasks successfully" (Linkenauger et al., 2015, p. 399), or "We found that minimal experience reaching with the virtual arm can influence perceived distance" (Linkenauger et al., 2015, p. 393), are used quite frequently. Sometimes this is undoubtedly the case, but it is by no means the rule as the results of the current studies highlight. Perhaps the present studies represent a unique situation where participants were acting with a full avatar rendered in VR which caused calibration to occur relatively slowly.

The results of the current research are mostly in accord with the main findings of van Andel et al. (2017) regarding calibration. The result of their literature review revealed that the timeframe for calibration to occur is variable (for a discussion of rate of calibration see Ebrahimi, Altenhoff, Pagano, & Babu, 2015; van Andel et al., 2017), and it is contingent upon the aptness of the information explained for calibration, and when the movement itself is explored, calibration occurs relatively quickly (van Andel et al., 2017). Perhaps by occluding participants' view of their avatar during their initial reach during the calibration stage this caused the process of calibration to occur relatively slowly. However, after the initial reach was made all participants were able to readjust their reach to be 100% accurate and this readjustment was made in full view, so participants were allowed to explore and perceive the results of their reaching movements in some way in the present studies.

The current work also relates to the second point raised by van Andel et al. (2017), namely that an investigation into individual differences in the process of calibration must be carried out. In repeated measures multilevel modeling, random slopes test to see if there are individual differences across L1 variables. For example, if there was a significant random slope for trial number this would indicated that there were individual differences in the rate of calibration. In Experiment 1, there was a significant random slope for trial number when predicting absolute error, suggesting that there are individual differences in the rate of calibration. This is certainly a topic for further study.

Important issues raised by current work

The current studies highlight two important issues. First, that the rate of calibration most likely differs across situations, and second, there is a lack of necessary criteria that exist in the literature to help define if calibration has occurred or not. In the given definitions of calibration there is no set criteria for what calibrated action looks like other than to produce environmentally directed action that is informed by a scaling between action and perception (Mon-Williams & Bingham, 2007). Generally, calibration is measured by or judged by the action judgments that are produced in a posttest after calibration has supposedly occurred. For example, if given a tool that increases reaching distance and participants calibrate to that tool, evidence for this is taken in the form of those participants perceiving further distances to be within reach in a posttest even after the tool has been removed from the system. Yet this way of measuring calibration does not investigate the actual process of calibration itself as it is occurring (see Bingham and Romack, 1999). As revealed in Experiment 1 in the current studies, in the calibration phase only, participants in the altered avatar condition consistently exhibited greater amounts of absolute error than participants whose action capabilities had not been manipulated. This same pattern of results, where calibration is said to have occurred but participants still exhibit error in their behavior, is seen in previous work as well (Kelly et al., 2013; Kelly et al., 2014; Mon-Williams & Bingham, 2007; Scott & Gray, 2010). Even in studies where the authors report that calibration occurred relatively quickly, errors in behavior compared to control groups are still evident. Can we confidently claim that calibration has occurred when the action judgments between two groups are similar, but there are differences between the groups in the error exhibited when carrying out the actual motor behaviors? Moving forward, it is important to consider both action judgments, defined as choosing to engage in a behavior or not (i.e. reaching to a target or not reaching to a target), and movement control, defined as level of accuracy or error in a completed action, as criteria for determining if calibration has occurred and how successful the process of calibration was. Many previous investigations into the process of calibration have ignored data produced in the calibration phase, and only relied on action judgments in the posttest to make claims that calibration has occurred. However, action judgments are not necessarily correlated with the actual accuracy of the movement control when carrying out an action. We believe it is very important to understand the control of movements in the calibration phase, perhaps defined as exhibited error, when attempting to show that calibration has occurred.

This is not to say that calibration has not occurred in any of these situations, but rather that investigators need to consider setting applicable criteria to help define how successful the process of calibration was in their studies. For example, in Experiment 1, it would have been ideal for participants in the altered avatar condition to demonstrate absolute error that was equivalent to the absolute error demonstrated by the normal avatar group or for those participants to have rescaled their reaches to an amount equal to the rescaling of the altered avatar arm. As stated previously, any behavior will involve noise so expecting calibration to result in perfect performance is not realistic. Rather, having calibration result in error terms produced by the behaviors that are largely similar to error terms exhibited by the control group is a tenable criterion. The same can be said for the commonly accepted criterion of action judgments, in that both group should make judgments based off the relationship between their action capabilities and the environment. Overall, we believe that both action judgments and movement control must be considered as necessary criteria to judge if calibration has actually occurred or not.

Another behavioral measure that could be used to address if calibration has occurred is precision. A limitation of the current analyses is that only accuracy of movements in reference to the target was considered. However, accuracy is independent of the precision of a movement. For instance, the consistency of reaches in each phase of the experiments could be quantified as well. Future work should investigate whether calibration affects both accuracy *and* precision. By creating an error term and a precision term, future analyses could investigate the effect that calibration has on accuracy and precision independently. More specifically, future work could also investigate the rate at which calibration affects accuracy, as done in the current studies, *and* the rate at which calibration affects precision of movements.

Moving on, one of the primary research questions of the current work was whether or not using an avatar with an extended arm in VR is akin to using a tool in the real world. The answer to this question is still somewhat unclear. In one regard, there are similarities between acting with a tool and an avatar because participants are able to calibrate to the extension of their reaching capabilities and this calibration is somewhat enduring. Numerous studies have shown that near space is perceived differently than far space and that manipulations to action capabilities can influence the perception of near space, meaning what is within reach (Berti and Frassinetti, 2000; Iriki, 1996, Witt et al., 2005, Witt, 2011). This same pattern occurred in Experiment 1 and Experiment 2. Participants in the altered avatar group made more incorrect judgments in the posttest than participants in the normal avatar group. That is, the participants in the altered avatar condition were more likely to either reach to targets that were unreachable or fail to reach to targets that were within reach. This finding also suggests that calibration to an altered avatar in the intervening calibration phase carried over to the posttest, in that participants in the altered avatar condition continued to reach to target distances that would have been reachable in the calibration phase with the altered avatar but were no longer reachable in the posttest.

Conversely, the results of the present work suggest that calibration to an altered avatar occurs more slowly than calibration to a handheld tool as identified in previous literature. Future research should test to see if calibration to anything, be it a tool or an extended arm, in VR in general takes longer than calibration in the real world. For example, a future experiment could test the rate of calibration to a handheld tool rendered in VR to the rate of calibration to an extended arm rendered in VR. Both of these conditions could then be compared to the rate at which an actor can calibrate to a handheld tool in the real world.

Discussion of Reversion

In addition to contributing to the literature on the rate at which calibration occurs, another unique contribution of the present research is the demonstration of reversion (as differentiated from calibration). We have defined reversion as a shift away from prior calibration to act in accordance with a stored body schema when no feedback is present. We believed reversion would be evidenced by participants in the altered avatar condition exhibiting the same amount of absolute error as participants in the normal avatar condition in the posttest. In other words, evidence of the use of a stored body schema would be demonstrated by participants in the altered avatar condition reverting back to reaching as though their arm was once again its normal length after feedback is removed. Nonetheless, participants in both conditions exhibited equal amounts of absolute error towards the end of the posttest in Experiment 1, meaning that both reached as though the arm was its normal length. The rate of calibration was found to be quicker in the calibration phase than the rate of reversion in the posttest when all reaching was performed in VR. However, no difference in the rate of calibration and reversion was found when the calibration occurred in VR and the revision in the real world. Perhaps this departure from what was hypothesized was affected by the switch of modalities, whereas all phases occurred in VR in Experiment 1, but in Experiment 2 the calibration phase occurred in VR and the posttest occurred in the real world.

One explanation for this finding regarding reversion is that every participant reverted back to acting in accord with a stored body schema based off their normal capabilities. In both Experiments participants in the normal avatar condition exhibited absolute error that was relatively constant across all three phases (regardless if there was feedback or not), indicating that their behaviors were constant and predictable. Thus, the reaching behaviors exhibited by participants in the altered avatar condition in the posttest in Experiment 1 after the removal of their extended avatar arm can be thought of as reversion to a stored body schema.

There is a similar finding in many of the prism goggle experiments, as discussed in the introduction, where it usually takes fewer trials to recalibrate in the final phase than it does to calibrate to the prism goggles in the exposure phase. The major difference between this finding stemming from the prism goggle literature and the current work is that participants received feedback in the final phase of prism goggle experiments (i.e. they could see the movement and result of their actions) whereas no feedback was available to participants in the posttest in the current research. Perhaps the reason for this increased rate of recalibration in the final phase of prism goggle experiments is because the recalibration it is accelerated by reversion back to a stored body schema. However, reversion is not necessarily an impediment to calibrating to altered action capabilities or sensory units when explicit informative feedback is available. Perhaps the process of (re)calibration works differently under different conditions. When acting with altered action capabilities and informative feedback is available, reversion does not inhibit or affect the process of calibration. When no feedback is available while experiencing a shift in action capabilities, calibration cannot be expected to occur. After experiencing a change to action capabilities, and returning to acting with relatively normal (meaning well known) action capabilities (meaning acting with our normal body in an everyday state) and feedback is available, reversion may even accelerate this recalibration, as shown in the prism goggle work. Then, as shown in the current work, reversion seems to appear after returning to regular action capabilities (after experiencing a change to action capabilities of course) and no feedback is available. This relationship could be represented in a two by two matrix of action capabilities (altered vs. regular) and feedback (available vs. not available).

One note must be made before continuing. In everyday life, there is always feedback of some sort available to inform an actor about the outcome of their behaviors in the world. Perhaps reversion only occurs under very artificial situations where there is no feedback available. Further support for the notion of reversion could be obtained by performing an experiment similar to the experiments performed in this paper. All participants would complete a pretest with no feedback available, then they would complete two calibration phases with feedback available. In the first calibration phase, participants would be tasked with calibrating to an avatar with an extended arm. Then in the second calibration phase, they would be tasked with calibrating to an avatar that faithfully represents their real body dimensions. If it is shown that calibration in the second calibration phase occurs more quickly than in the first calibration phase, further evidence for reversion as an accelerating agent of calibration would be obtained. In a sense, this proposed study could be treated as the virtual replication of the traditional prism goggle studies.

Another possible experiment would involve all participants performing a pretest, and then calibrating to an altered avatar. Next, half of the participants would perform a posttest without feedback, and the other half would receive a second calibration phase with a normal arm. Comparison of the third phase would show if reversion is different than recalibration.

Implications for the Body Schema

The present results have implications for conceptualizations of the body schema, namely that the body schema is both malleable and stable at the same time. Generally, research into the process of calibration has highlighted the plasticity of the human perception-action system in responding to discrepancies (Bingham & Romack, 1999; Bingham & Pagano, 1998; Mon-Williams & Bingham, 2007). Just as the process of calibration works to keep behavior constant across perturbations, such as sensory or action based perturbations, the body schema is similar. To some extent there must be a stable body schema that is not subject to changes over short timescales or relatively minor perturbations. However, it would not be functionally efficient for the body schema to be permanent, as it is a fact that our action capabilities change on a regular basis, such when we use tools or across the lifespan due to changes like increases in strength and acquiring new skills.

Crucial aspects of the body schema may be perceived on-line. In this sense, the body schema is fluid, and perceived continuously as the limbs and their attachments change (Maravita & Iriki, 2004; Pagano & Turvey, 1998). A key finding from previous literature is that both limbs and hand-held objects are perceived through the same mechanism. That is, the same principles underlie both the perception of attached objects and the perception of the body itself (Pagano & Turvey, 1998). This finding links our understanding regarding the malleability of the body schema and our understand of how attached objects are perceived and then incorporated into the body schema, because our perception-action system treats them like they are part of the body. The body schema does not seem to distinguish between objects and limbs, but rather it represents the effects of a tool as the lengthening of the arm that is incorporated into the body schema (Cardinali et al., 2009; Maravita & Iriki, 2004; Sposito et al., 2012). Perhaps calibration is what allows the body schema to provide the means by which the perception-action system maintains constant behavior across perturbations.

One goal of the present work was to extend and further test these ideas by investigating the malleability of the on-line body schema in the context of reaching in VR with an extended avatar arm. While a temporary 'online' body scheme is altered by calibration, a more permanent stored body schema likely exists simultaneously. Experiences that cause the more permeant body schema to be altered should be a topic for further study (though many have already talked about this).

The present work also has practical ramifications as well. For example, this work has implications for accepting a limb that is bigger, or smaller, than your own limb, such as when amputees receive artificial limbs (Imaizumi, Asai, & Koyama, 2016). These artificial limbs may or may not be the exact same size as their lost limb, and they likely do not possess the same weight properties either, as artificial limbs are often lightweight. The results of the present work suggest that calibration to these altered limbs is possible, but only after numerous experiences using them, and that the artificial limbs can be incorporated into the body schema through the process of calibration.

Comparison of Reaching in the Real World to Virtual Reality

The direct comparison of blind reaching in VR to blind reaching in the real world revealed that reaching with an avatar in VR resulted in less absolute error than reaching in the real world. Participants in Experiment 2 (acting in the real world) demonstrated almost two and a half centimeters more error on average than participants in Experiment 1 (in the normal arm length condition). Based on the current finding, it seems that acting with an avatar in VR gives an added benefit of decreased disparity between target distance and estimated distance when reaching to targets in the virtual world compared to the real world.

Previous research has pointed out that virtual environments are perceived differently than normal environments, and that distance perception is especially affected in virtual environments (Bingham et al., 2001; Ebrahimi et al., 2016; Mon-Williams & Bingham, 2007; Napieralski et al., 2011). Previous investigators have concluded that results obtained in virtual environments may not be representative of how people behave in the real world. Other work has found that the presence of an avatar serves to alleviate differences between perception-action in virtual environments and the real world (Mohler et al., 2010).

For example, previous research has demonstrated that providing an avatar allows for people to act more similarly to how they would in the real world than when an avatar is not provided. Lin, Reiser and Bodenheimer (2015) found that providing a self-avatar in a virtual environment generates action judgments that are not significantly different from action judgments made in the real world. The current research adds to this finding by demonstrating that avatars that faithfully represent the anthropometric dimensions of a user allow users to behave in the virtual environment in a manner that is most similar to real world behaviors immediately. However, if an avatar is provided but the avatar possesses different dimensions than your real-world body, exposure to using that avatar by acting with it and receiving feedback about your performance is necessary for calibration to occur. In terms of correct judgments, the present work shows that switching from acting with an altered avatar to a faithful avatar results in action judgments that are significantly different than the decisions exhibited by the normal avatar group. This suggests that the avatar a user is acting with cannot change while they are using it for their action judgments to remain similar to how they would act in the real world.

Future Research

Many important questions remain in regards to investigating the process of calibration to an avatar in VR. As identified by van Andel et al. (2017), very little work exists investigating calibration in older adults. The author strongly believes that a study comparing the rate of calibration between younger and older adults is necessary in order to come to a complete understanding of how calibration works across the lifespan.

One notable aspect missing from the current work is the lack of a baseline condition to compare all other conditions to. Future work should include two more groups of participants who perform the exact same tasks as in the current research, except all reaches will be completed in the real world. One group of participants will reach with a tool that does not functionally increase their reach for all three phases. The other group will reach with a tool that does not functionally increase their reach in the pretest and posttest, but will be given a long tool in the calibration phase. Performance on the reaching task, and rate of calibration can then be compared amongst all three experiments (VR only, mix of RW and VR, and RW only).

Next, we were able to demonstrate that calibration is not contingent upon the distances presented during training, but that calibration extends to the full range of reachable distances in near space. However, future work should investigate whether the rescaling of perceived space that is coupled with calibrating to an avatar with extended anthropometric dimensions generalizes to distances that are out of reach. For example, does the rescaling evidenced in the present work also generalize to a rescaling of perceived far space (i.e. that space that is surely out of reach)?

Another study that should be conducted to assess further the need for VR designers to scale avatars to individual users would only make use of one avatar. Each participant recruited for the study would be treated as their own independent variable, as they would all possess bodies of different anthropometric dimensions in the real world. In this way, the question of whether individual differences in body sizes affects calibration to an avatar of one standard size could be investigated. The results of this investigation would have direct implications for the need of VR designers to be able to provide every individual user with a properly scaled avatar, or if individual users all have the ability to calibrate to the action capabilities of one standard size avatar while acting in VR.

Another question that deserves study relates to previous work completed on visual capture of felt body position. A major finding from this body of literature is that when there is a conflict between the proprioceptive position of one's arm and the visually specified position of one's arm, people tend to resolve this conflict by relying on the visual position. The resulting experience is that people feel their arm to be where it is seen (see Slater et al., 2008). This prior work has a significant relation to the current work. In the present studies participants never dealt with a conflict between the felt position of their end effector and the visually specified position of their avatar arm as each avatar type was matched with a tool that extended participants' real world reaches to be equivalent to the extension of reach seen in VR. Future studies should manipulate the length of the tool that participants are holding in the real world to create a conflict between the visually specified virtual arm and the felt position of the end effector in the real world. This study would involve both a normal and altered avatar, as well as a congruent (i.e. no conflict between seen and felt position) and a non-congruent (i.e. a conflict between seen and felt positon) group. A test of the visual dominance theory as it applies to virtual reality could be conducted in this way, and one could hypothesize that there would be significant differences in distance estimations and the rate of calibration between the congruent and non-congruent groups for both types of avatar. The authors believe that novel hypotheses, findings, and interpretations would reveal themselves in a test of the commonly accepted visual dominance theory.

Similarly, future studies could extend the current work with a condition where the participant only uses their hand to act in the real world, regardless of the length of the avatar arm presented to them in VR. Incorporating this condition would also create a mismatch between the visually specified position of the avatar arm in VR and the felt position of the hand. In this scenario, one would expect for the visually specified information to become dominant over the felt position of the hand due to the visual dominance phenomenon.

There are many other research questions and studies that need to be conducted stemming from the present work. I believe it would be important to identify the rate of calibration in a more realistic setting. This could be accomplished by replicating the current work, but providing experiments both visual and tactile feedback in each phase (i.e. make each phase closed loop). Also, as it relates to a guiding question of the current work, a future question that needs to be answered regards reaching with a virtually rendered tool compared to reaching with a virtually rendered avatar arm to explore if there are any differences in calibration and reaching behaviors between the two conditions. Another interesting question would investigate the effect that a time delay between the calibration phase and posttest has on reaching behaviors. In the current work, steps were taken to ensure that the posttest occurred as quickly after the calibration phase as possible. For example, future studies could manipulate the delay between these two phases (i.e. immediate, one minute, five minutes, and ten minutes) to test how long calibration remains engrained in the system and / or if the process of reversion is time dependent. Lastly, one other study should consider the effect of using an avatar on sense of presence and ownership in VR. Participants could be administered the body ownership questionnaire after the familiarization phase and then again after the experimental calibration phase as well. This type of study could investigate could compare how the responses change over the course of the experiment acting with different types of avatar.

Applications of Current Work

Both of the present studies revealed interesting perception-action mechanisms regarding the process of calibration. Both studies, especially Experiment 2, have very important practical implications as well. VR technology has numerous current applications, and the number of meaningful applications is growing. Currently, VR is being used to aid rehabilitation, as a training exercise in many fields (including but not limited to the medical field and combat), education, behavioral research, and entertainment. Most technology companies offer some sort of virtual reality gaming product. For any of these uses of VR technology to be maximally useful, especially when used as a training aid, a user's perception-action processes in VR must match their perception-action processes in the real world.

Previous research has highlighted the importance of providing a user with an avatar when acting in VR (Mohler et al., 2010). However, as demonstrated in the current work and other related work (Linkenauger et al., 2015), merely providing an avatar does not tell the whole story. There are major implications for how a user will perceive and act in the immersive virtual environment based off the size and dimensions of their avatar. The present results support the idea that calibration to an avatar that possesses different anthropometric dimensions than your own body is possible, but this process is relatively lengthy in comparison to other reported rates of calibration in that it takes approximately 45 trials over the span of 10 to 15 minutes to occur. Further, participants who were given an avatar with extended arms produced behaviors that involved more discrepancies between the presented target distance and their estimated reach distance over the course of the training phase than those participants who reached with an avatar scaled to the size of their body. This finding, based off data collected in the calibration phase and posttest, suggests that for training in VR to be most effective, the avatar given to a user must faithfully represent the dimensions of their body.

Of note, within the current studies, participants performed their reaching behaviors in VR or the real world over the course of about an hour. Thus, the current findings can only generalize to applications where VR is used for training during this time frame. Future research should investigate if repeated exposures (i.e. multiple hours of training over multiple days) to VR training with an altered avatar results in similar findings.

The major finding derived from comparing data produced in the posttest only in Experiment 1 and in Experiment 2 also has important connotations for applications of VR used as a training aid. In the posttest only, participants who completed the posttest in the real world exhibited greater absolute error in their reaches than participants who completed the posttest in VR. This result suggests that there seems to be a negative effect on reaching to target distances after switching modalities. Participants who completed the calibration phase in VR and the posttest in the real world (i.e. participants in Experiment 2) exhibited greater disparities in reaching to targets. Those participants who acted in a congruent modality for each phase (in VR for both phases) showed significantly less disparity in their reaches to targets. More specifically, in both experiments, participants in the altered avatar group exhibited larger amounts of absolute error in the posttest than participants in the normal avatar condition.

The findings of the current work support the notion that for training conducted in VR to be maximally effective in the real world the size of the avatar must faithfully represent the user. Further, training received in VR, with an avatar arm that faithfully represents the size of the actor's arm, that is then applied in the real world resulted in reaching behaviors that involved less disparity between target distances and estimated distances in the real world than before training occurred. But training in VR with an avatar arm that is different from one's own arm resulted in greater disparities between target distances and estimated distances in the real world after training occurred. There is an issue of transfer across modalities from training received in VR to practice in the real world. This is a topic that deserves further attention as well. Overall, the findings from these studies have direct implications for how avatars are designed and presented to the user in immersive virtual environments. Presenting users with avatars that do not represent their normal body cause a change in how the virtual environment is perceived, and cannot be used to act in a manner that is representative of how that user would act with an avatar that is designed to match their bodily dimensions. Based on the current findings, it is imperative that VR developers take the necessary steps to ensure that users are presented with and can act with an avatar that faithfully matches the dimensions of their body if the virtual environment is to be used for training that must translate back to the real world. If transfer of training is not a key concern, the results support the conclusion that participants are in fact able to calibrate effectively to an avatar that possesses longer arms than their own body.

APPENDICES

Appendix A

Body Ownership Questionnaire

Body Ownership

1. When you were looking down from above how much did you feel a strong connection with the avatar as if you were looking down at yourself?

NOT AT ALL 0	1 2 3 4 5 6 7 8 9	10 VERY MUCH

2. How much did you feel that the seated avatar's body was your body?

NOT AT ALL 0	1 2 3 4 5 6 7 8 9	10 VERY MUCH

3. How strong was the feeling that the movements of the avatar were caused by your own movements?

NOT AT ALL 0	1 2 3 4 5 6 7 8 9	10 VERY MUCH

4. How much did you feel that the virtual body was another person?

NOT AT ALL 0	1 2 3 4 5 6 7 8 9	10 VERY MUCH

5. How much was this experience more like watching a scene from the outside compared to being part of the scene?

NOT AT ALL 0	1 2 3 4 5 6 7 8 9	10 VERY MUCH

6. How strong was the feeling that the body of the person in the mirror was your body?

NOT AT ALL 0	1 2 3 4 5 6 7 8 9	10 VERY MUCH

REFERENCES

Altenhoff, B.M., Napieralski, P.E., Long, L.O., Bertrand, J.W., Pagano, C.C., Babu, S.V., & Davis, T.A. (2012). Effects of Visual and Haptic Feedback on Near-Field Depth Perception in an Immersive Virtual Environment. *Proceedings of the ACM Symposium on Applied Perception, Los Angeles*, CA, Aug 3-4, 2012.

Banakou, D., Groten, R., & Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, *110*(31), 12846-12851.

Berti, A., & Frassinetti, F. (2000). When far becomes near: Remapping of space by tool use. *Journal of Cognitive Neuroscience*, *12*(3), 415-420.

Bertram, J., Moskaliuk, J., & Cress, U. (2015). Virtual training: Making reality work? *Computers in Human Behavior*, *43*, 284-292.

Bingham, G. P., Bradley, A., Bailey, M., & Vinner, R. (2001). Accommodation, occlusion, and disparity matching are used to guide reaching: a comparison of actual versus virtual environments. *Journal of experimental psychology: human perception and performance*, *27*(6), 1314-1334.

Bingham, G. P., & Pagano, C. C. (1998). The necessity of a perception–action approach to definite distance perception: Monocular distance perception to guide reaching. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 145.

Bingham, G. P., Pan, J. S., & Mon-Williams, M. A. (2014). Calibration is both functional and anatomical. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 61-70.

Bingham, G. P., & Romack, J. L. (1999). The rate of adaptation to displacement prisms remains constant despite acquisition of rapid calibration. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1331-1346.

Bliese, P. D. (1998). Group size, ICC values, and group-level correlations: A simulation. *Organizational Research Methods*, *1*(4), 355-373.

Bliss, J. P., Tidwell, P. D., & Guest, M. A. (1997). The effectiveness of virtual reality for administering spatial navigation training to firefighters. *Presence: Teleoperators and Virtual Environments*, *6*(1), 73-86.

Bongers, R. M., Michaels, C. F., & Smitsman, A. W. (2004). Variations of tool and task characteristics reveal that tool-user postures are anticipated. *Journal of Motor Behavior*, 36, 305-315.

Bourgeois, J., & Coello, Y. (2012). Effect of visuomotor calibration and uncertainty on the perception of peripersonal space. *Attention, Perception, & Psychophysics*, *74*(6), 1268-1283.

Bufton, A., Campbell, A., Howie, E., & Straker, L. (2014). A comparison of the upper limb movement kinematics utilized by children playing virtual and real table tennis. *Human Movement Science*, *38*, 84-93.

Burton, G. (1993). Non-neural extensions of haptic sensitivity. *Ecological Psychology, 5*, 105-124.

Carello, C., Fitzpatrick, P., & Turvey, M. T. (1992). Haptic probing: Perceiving the length of a probe and the distance of a surface probed. *Perception & Psychophysics*, *51*(6), 580-598.

Cardinali, L. Frassinetti, F., Brozzoli, C., Urquizar, C., Roy, A. C., & Farne, A. (2009). Tool-use induces morphological updating of the body schema. *Current Biology, 19*, R478-R479.

Coats, R. O., Pan, J. S., & Bingham, G. P. (2014). Perturbation of perceptual units reveals dominance hierarchy in cross calibration. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(1), 328-341.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied multiple correlation/regression analysis for the behavioral sciences. *UK: Taylor & Francis*.

Chemero, A. (2003). An outline of a theory of affordances. *Ecological psychology*, *15*(2), 181-195.

Creem-Regehr, S. H., Stefanucci, J. K., & Thompson, W. B. (2015). Perceiving Absolute Scale in Virtual Enviornments: How Theory and Application Have Mutually Informed the Role of Body-Based Perception. In *Psychology of Learning and Motivation,* Brian H. Ross (ed.). Academic Press, 195-224.

Ebrahimi, E., Altenhoff, B., Hartman, L., Jones, J. A., Babu, S. V., Pagano, C. C., & Davis, T. A. (2014, August). Effects of visual and proprioceptive information in visuo-motor calibration during a closed-loop physical reach task in immersive virtual environments. In *Proceedings of the ACM Symposium on Applied Perception* (ACM SAP), 103-110, Vancouver, BC.

Ebrahimi, E., Altenhoff, B. M., Pagano, C. C., & Babu, S. V. (2015, March). Carryover effects of calibration to visual and proprioceptive information on near field distance judgments in 3D user interaction. In *2015 IEEE Symposium on 3D User Interfaces* (3DUI), 97-104), Arles, France.

Ebrahimi, E., Babu, S. V., Pagano, C. C., & Jörg, S. (2016). An empirical evaluation of visuo-haptic feedback on physical reaching behaviors during 3D interaction in real and immersive virtual environments. *ACM Transactions on Applied Perception* (TAP), 13, 19:1-19:21.

Fajen, B. R. (2005). Perceiving possibilities for action: On the necessity of calibration and perceptual learning for the visual guidance of action. *Perception*, *34*(6), 717-740.

Fajen, B. R. (2007). Rapid recalibration based on optic flow in visually guided action. *Experimental Brain Research*, *183*(1), 61-74.

Fajen, B. R., Riley, M. A., & Turvey, M. T. (2009). Information, affordances, and the control of action in sport. *International Journal of Sport Psychology*, *40*(1), 79-107.

Fitzpatrick, P., Carello, C., & Turvey, M. T. (1994). Eigenvalues of the inertia tensor and exteroception by the "muscular sense*." Neuroscience, 60*, 551-568.

Franchak, J. M., van der Zalm, D. J., & Adolph, K. E. (2010). Learning by doing: Action performance facilitates affordance perception. *Vision Research*, *50*(24), 2758-2765.

Ganier, F., Hoareau, C., & Tisseau, J. (2014). Evaluation of procedural learning transfer from a virtual environment to a real situation: a case study on tank maintenance training. *Ergonomics*, *57*(6), 828-843.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston, MA: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Grabowski, A., & Jankowski, J. (2015). Virtual Reality-based pilot training for underground coal miners. *Safety Science*, *72*, 310-314.

Head. H. (1920). *Studies in neurology, Vol. 2*. London: Oxford University Press.

Head, H. & Holmes, G. (1911-1912). Sensory disturbances from cerebral lesions. *Brain, 34*, 102-254.

Heft, H. (1993). A methodological note on overestimates of reaching distance: Distinguishing between perceptual and analytical judgments. *Ecological Psychology*, *5*(3), 255-271.

Heft, H. (2017). Perceptual Information of "An Entirely Different Order": The "Cultural Environment" in The Senses Considered as Perceptual Systems. *Ecological Psychology*, *29*(2), 122-145.

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications*. Routledge.

Hyltander, A., Liljegren, E., Rhodin, P. H., & Lönroth, H. (2002). The transfer of basic skills learned in a laparoscopic simulator to the operating room. *Surgical Endoscopy and Other Interventional Techniques*, *16*(9), 1324-1328.

Imaizumi, S., Asai, T., & Koyama, S. (2016). Embodied prosthetic arm stabilizes body posture, while unembodied one perturbs it. *Consciousness and Cognition*, *45*, 75-88.

Iodice, P., Scuderi, N., Saggini, R., & Pezzulo, G. (2015). Multiple timescales of body schema reorganization due to plastic surgery. *Human Movement Science*, *42*, 54-70.

Jun, E., Stefanucci, J. K., Creem-Regehr, S. H., Geuss, M. N., & Thompson, W. B. (2015). Big foot: Using the size of a virtual foot to scale gap width. *ACM Transactions on Applied Perception (TAP)*, *12*(4), 16.

Kelly, J. W., Donaldson, L. S., Sjolund, L. A., & Freiberg, J. B. (2013). More than just perception–action recalibration: Walking through a virtual environment causes rescaling of perceived space. *Attention, Perception, & Psychophysics*, *75*(7), 1473-1485.

Kelly, J. W., Hammel, W. W., Siegel, Z. D., & Sjolund, L. A. (2014). Recalibration of perceived distance in virtual environments occurs rapidly and transfers asymmetrically across scale. *IEEE transactions on visualization and computer graphics*, *20*(4), 588-595.

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, *21*(4), 373-387.

Kilteni, K., Normand, J. M., Sanchez-Vives, M. V., & Slater, M. (2012). Extending body space in immersive virtual reality: a very long arm illusion. *PloS One*, *7*(7), e40867.

Kokkinara, E., Slater, M., & López-Moliner, J. (2015). The effects of visuomotor calibration to the perceived space and body, through embodiment in immersive virtual reality. *ACM Transactions on Applied Perception (TAP)*, *13*(1), 3.

Koritnik, T., Koenig, A., Bajd, T., Riener, R., & Munih, M. (2010). Comparison of visual and haptic feedback during training of lower extremities. *Gait & posture*, *32*(4), 540-546.

Kozak, J. J., Hancock, P. A., Arthur, E. J., & Chrysler, S. T. (1993). Transfer of training from virtual reality. *Ergonomics*, *36*(7), 777-784.

Larrue, F., Sauzeon, H., Wallet, G., Foloppe, D., Cazalets, J. R., Gross, C., & N'Kaoua, B. (2014). Influence of body-centered information on the transfer of spatial learning from a virtual to a real environment. *Journal of Cognitive Psychology*, *26*(8), 906-918.

Lessard, D. A., Linkenauger, S. A., & Proffitt, D. R. (2009). Look before you leap: Jumping ability affects distance perception. *Perception*, *38*(12), 1863-1866.

Leyrer, M., Linkenauger, S. A., Bülthoff, H. H., Kloos, U., & Mohler, B. (2011, August). The influence of eye height and avatars on egocentric distance estimates in immersive virtual environments. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, 67-74, ACM.

Lin, Q., Rieser, J., & Bodenheimer, B. (2012, August). Stepping over and ducking under: The influence of an avatar on locomotion in an HMD-based immersive virtual environment. In *Proceedings of the ACM Symposium on Applied Perception* (pp. 7-10). ACM.

Lin, Q., Rieser, J. J., & Bodenheimer, B. (2013, August). Stepping off a ledge in an HMD-based immersive virtual environment. In *Proceedings of the ACM Symposium on Applied Perception* (pp. 107-110). ACM.

Lin, Q., Rieser, J., & Bodenheimer, B. (2015). Affordance judgments in HMD-based virtual environments: Stepping over a pole and stepping off a ledge. *ACM Transactions on Applied Perception (TAP)*, *12*(2), 6.

Linkenauger, S. A., Bülthoff, H. H., & Mohler, B. J. (2015). Virtual arm′ s reach influences perceived distances but only after experience reaching. *Neuropsychologia*, *70*, 393-401.

Lintern, G., Roscoe, S. N., Koonce, J. M., & Segal, L. D. (1990). Transfer of landing skills in beginning flight training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *32*(3), 319-327.

Lotze, R. H. 1885. *Mikrokosmus* (Vol. 1, 4th ed.). Translated by E. Hamilton and E. E. C. Jones. Edinburgh: Clark (Original work published 1856).

Lotze, R. H. 1973. *Outlines of Psychology*. Translated by C. L. Herrick. Minneapolis, MN: Williams (Original work published 1885).

Maravita, A. & Iriki, A. (2004). Tools for the body (schema). *TRENDS in Cognitive Sciences*, *8*(2).

Mark, L. S. (1987). Eyeheight-scaled information about affordances: a study of sitting and stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 361.

Mark, L. S., Balliett, J. A., Craver, K. D., Douglas, Stephen D., and Fox, T. (1990). What an actor must do in order to perceive the affordance for sitting. *Ecological Psychology, 2*(4), 325-366.

Maselli, A., & Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience*, *7*, 1-15.

McManus, E. A., Bodenheimer, B., Streuber, S., De La Rosa, S., Bülthoff, H. H., & Mohler, B. J. (2011, August). The influence of avatar (self and character) animations on distance estimation, object interaction and locomotion in immersive virtual environments. In *Proceedings of the ACM SIGGRAPH Symposium on applied perception in graphics and visualization* (pp. 37-44). ACM.

Merleau-Ponty, M. (1962). Phenomenology of perception, trans. Colin Smith.

Mohler, B. J., Creem-Regehr, S. H., Thompson, W. B., & Bülthoff, H. H. (2010). The effect of viewing a self-avatar on distance judgments in an HMD-based virtual environment. *Presence: Teleoperators and Virtual Environments*, *19*(3), 230-242.

Mon-Williams, M., & Bingham, G. P. (2007). Calibrating reach distance to visual targets. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(3), 645-656.

Napieralski, P.E., Altenhoff, B.M., Bertrand, J.W., Long, L.O., Babu, S.V., Pagano, C.C., Kern, J. & Davis, T.A. (2011). Comparing Near Field Distance Perception in Real and Virtual Environments Using Both Verbal and Action Responses. *ACM Transactions on Applied Perception*, *8*(3), 18:1-18:19.

Pagano, C. C., & Turvey, M. T. (1992). Eigenvectors of the inertia tensor and perceiving the orientation of a hand-held object by dynamic touch. *Perception & Psychophysics*, *52*(6), 617-624.

Pagano, C. C., & Turvey, M. T. (1993). Perceiving by dynamic touch the distances reachable with irregular objects. *Ecological psychology*, *5*(2), 125-151.

Pagano, C.C., & Turvey, M. T. (1995). The inertia tensor as a basis for the perception of limb orientation. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1070-1087.

Pagano, C. C., & Turvey, M. T. (1998). Eigenvectors of the inertia tensor and perceiving the orientation of limbs and objects. *Journal of Applied Biomechanics, 14*, 331-359.

Proffitt, D. R., & Linkenauger, S. A. (2013). Perception viewed as a phenotypic expression. *Action Science: Foundations of an Emerging Discipline*, *171-197*.

Ramenzoni, V. C., Davis, T. J., Riley, M. A., & Shockley, K. (2010). Perceiving action boundaries: learning effects in perceiving maximum jumping-reach affordances. *Attention, Perception, & Psychophysics*, *72*(4), 1110-1119.

Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR)*, *46*(2), 23.

Regian, J. W. (1997). Virtual reality for training: Evaluating transfer. In *Virtual Reality, Training's Future?* Springer US, 31-40.

Ries, B., Interrante, V., Kaeding, M., & Phillips, L. (2009, November). Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology* (pp. 59-66). ACM.

Rose, F. D., Attree, E. A., Brooks, B. M., Parslow, D. M., & Penn, P. R. (2000). Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics*, *43*(4), 494-511.

Scott, S., & Gray, R. (2010). Switching tools: Perceptual-motor recalibration to weight changes. *Experimental brain research*, *201*(2), 177-189.

Slater, M., Pérez-Marcos, D., Ehrsson, H., & Sanchez-Vives, M. V. (2008). Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience*, *2*, 6.

Slater, M., Pérez-Marcos, D., Ehrsson, H., & Sanchez-Vives, M. V. (2009). Inducing illusory ownership of a virtual body. *Frontiers inNneuroscience*, *3*, 214-220.

Slater, M., Spanlang, B., Sanchez-Vives, M. V., & Blanke, O. (2010). First person experience of body transfer in virtual reality. *PloS one*, *5*(5), e10564.

Sposito, A., Bolognini, N., Vallar, G., Maravita, A. (2012). Extension of perceived arm length following tool-use: Clues to plasticity of body metrics. *Neuropsychologia, 50*, 2187–2194.

Stereo Fly Test. Stereo Optical, Chicago, IL.

Stoffregen, T. A., Yang, C. M., Giveans, M. R., Flanagan, M., & Bardy, B. G. (2009). Movement in the perception of an affordance for wheelchair locomotion. *Ecological Psychology*, *21*(1), 1-36.

Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, *4*(3), 173-187.

van Andel, S., Cole, M. H., & Pepping, G. J. (2017). A systematic review on perceptual-motor calibration to changes in action capabilities. *Human Movement Science*, *51*, 59-71.

Welch, R. (1986). Adaptation of space perception. *Handbook of Perception and Human Performance.*, *1*, 24, 1-45.

Welch, R. B., Bridgeman, B., Anand, S., & Browman, K. E. (1993). Alternating prism exposure causes dual adaptation and generalization to a novel displacement. *Perception & Psychophysics, 54,* 195-204.

Witt, J. K., Proffitt, D. R., & Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 880-888.

Wagman, J. B. (2012). Perception of maximum reaching height reflects impending changes in reaching ability and improvements transfer to unpracticed reaching tasks. *Experimental Brain Research*, *219*(4), 467-476.

Wagman, J. B., Caputo, S. E., & Stoffregen, T. A. (2016). Sensitivity to hierarchical relations among affordances in the assembly of asymmetric tools. *Experimental Brain Research*, 1-11.

Wagman, J. B., & Chemero, A. (2014). The end of the debate over extended cognition. In *Neuroscience, Neurophilosophy and Pragmatism* (pp. 105-124). Palgrave Macmillan UK.

Wagman, J. B., Taheny, C. A., & Higuchi, T. (2014). Improvements in perception of maximum reaching height transfer to increases or decreases in reaching ability. *The American Journal of Psychology*, *127*(3), 269-279.

Wagman, J. B., & Taylor, K. R. (2005a). Perceiving affordances for aperture crossing for the person-plus-object system. *Ecological Psychology, 17*, 105-130.

Wagman, J. B., Thomas, B. J., McBride, D. M., & Day, B. M. (2013). Perception of maximum reaching height when the means of reaching are no longer in view. *Ecological Psychology*, *25*(1), 63-80.

Warren Jr, W. H., & Whang, S. (1987). Visual guidance of walking through apertures: body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 371-383.

Willemsen, P., Gooch, A. A., Thompson, W. B., & Creem-Regehr, S. H. (2008). Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, *17*(1), 91-101.

Withagen, R., & Michaels, C. F. (2004). Transfer of calibration in length perception by dynamic touch. *Perception & Psychophysics*, *66*(8), 1282-1292.

Withagen, R., & Michaels, C. F. (2007). Transfer of calibration between length and sweet-spot perception by dynamic touch. *Ecological Psychology*, *19*(1), 1-19.

Witt, J. K., & Proffitt, D. R. (2008). Action-specific influences on distance perception: a role for motor simulation. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(6), 1479-1492.

Witt, J. K., Proffitt, D. R., & Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 880-888